

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2004, 21:50:48 ; Search time 1396 Seconds
(without alignments)
869.345 Million cell updates/sec

Title: US-10-028-415-2
Perfect score: 28
Sequence: 1 gctctgaactgcatccaaattcagggttc 28

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 1846492

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:
1: gb_ba:
2: gb_hg:
3: gb_in:
4: gb_om:
5: gb_ov:
6: gb_pat:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_sts:
12: gb_sy:
13: gb_un:
14: gb_vi:
15: em_ba:
16: em_fun:
17: em_hum:
18: em_in:
19: em_mu:
20: em_om:
21: em_or:
22: em_ov:
23: em_pat:
24: em_ph:
25: em_pl:
26: em_ro:
27: em_sts:
28: em_un:
29: em_vi:
30: em_hg_hum:
31: em_hg_inv:
32: em_hg_other:
33: em_hg_mus:
34: em_hg_pln:
35: em_hg_rod:
36: em_hg_mam:
37: em_hg_vrt:
38: em_sy:
39: em_hg_hum:
40: em_hg_mus:
41: em_hg_other:

Pred. No. is the number of results predicted by chance to have a

SUMMARIES

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	28	100.0	28	6	AR071972	AR071972 Sequence
2	28	100.0	28	6	AR071984	AR071984 Sequence
3	28	100.0	50	6	AR072005	AR072005 Sequence
4	16	57.1	25	6	AR071999	AR071999 Sequence
5	16	57.1	36	6	A62083	A62083 Sequence 18
6	16	57.1	36	6	A66040	A66040 Sequence 22
7	16	57.1	36	6	AR08019	AR08019 Sequence
8	16	57.1	36	6	AR154830	AR154830 Sequence
9	16	57.1	36	6	AR343590	AR343590 Sequence
10	15.4	55.0	20	6	AX4613512	AX4613512 Sequence
11	15.2	54.3	65	6	AX486149	AX486149 Sequence
12	15.2	54.3	82	6	AX386437	AX386437 Sequence
13	15.2	54.3	84	9	AF357525	AF357525 Homo sapi
14	15.2	54.3	99	6	AX769938	AX769938 Sequence
15	15.2	54.3	99	6	AX769969	AX769969 Sequence
16	15.2	54.3	99	6	AX769972	AX769972 Sequence
17	15	53.6	24	6	AX039922	AX039922 Sequence
18	15	53.6	28	6	BD248540	BD248540 Virulence
19	15	53.6	43	6	I15393	I15393 Sequence 15
20	15	53.6	68	6	A39690	A39690 Sequence 3
21	15	53.6	68	6	A39691	A39691 Sequence 4
22	15	53.6	68	6	AR074511	AR074511 Sequence
23	15	53.6	68	6	AR074512	AR074512 Sequence
24	15	53.6	100	9	HUMPCB01	M83365 Human plasm
25	14.8	52.9	54	6	AX794977	AX794977 Sequence
26	14.8	52.9	75	9	HSSNRPNS	AF087649 Homo sapi
27	14.8	52.9	87	8	AB027841	AB027841 Schizosac
28	14.6	52.1	59	6	I40084	I40084 Sequence 2
29	14.6	52.1	63	6	AX322070	AX322070 Sequence
30	14.6	52.1	63	6	I09547	I09547 Sequence 2
31	14.6	52.1	63	6	I09550	I09550 Sequence 5
32	14.6	52.1	65	6	AX485541	AX485541 Sequence
33	14.6	52.1	66	6	E06035	E06035 Partial seq
34	14.6	52.1	66	6	E06644	E06644 Synthetic D
35	14.6	52.1	66	6	E07704	E07704 DNA encodin
36	14.6	52.1	66	6	I40083	I40083 Sequence 1
37	14.6	52.1	66	6	AX322069	AX322069 Sequence
38	14.6	52.1	84	9	S41504	S41504 beta-globin
39	14.6	52.1	98	6	AX795637	AX795637 Sequence
40	14.6	52.1	98	6	AX796929	AX796929 Sequence
41	14.6	52.1	98	6	AX797191	AX797191 Sequence
42	14.6	52.1	100	6	AX795636	AX795636 Sequence
43	14.6	52.1	100	6	AX796928	AX796928 Sequence
44	14.6	52.1	100	6	AX797190	AX797190 Sequence
45	14.4	51.4	45	6	AX720715	AX720715 Sequence

ALIGNMENTS

RESULT 1
AR071972

LOCUS AR071972
DEFINITION Sequence 2 from patent US 5912168.
ACCESSION AR071972
VERSION AR071972.1 GI:7222860
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.

linear PAT 18-FEB-2000

REFERENCE 1 (bases 1 to 28)
AUTHORS Watson, J.D. and Rudert, F.
TITLE CD95 regulatory gene sequences
JOURNAL Patent: US 5912168-A 2 15-JUN-1999;
FEATURES Location/Qualifiers

```

source
1. .28
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match
Best Local Similarity 100.0%; Score 28; DB 6; Length 28;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCGGAAGTGCATCCAAATTCAGGTTTC 28
Db 1 GTCGGAAGTGCATCCAAATTCAGGTTTC 28

RESULT 2
AR071984/c
LOCUS AR071984 28 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 14 from patent US 5912168.
ACCESSION AR071984
VERSION AR071984.1 GI:7222872
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 28)
AUTHORS Watson,J.D. and Rudert,F.
TITLE CD95 regulatory gene sequences
JOURNAL Patent: US 5912168-A 14 15-JUN-1999;
FEATURES
Location/Qualifiers
source
1. .28
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match
Best Local Similarity 100.0%; Score 28; DB 6; Length 28;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCGGAAGTGCATCCAAATTCAGGTTTC 28
Db 28 GTCGGAAGTGCATCCAAATTCAGGTTTC 1

RESULT 3
AR072005
LOCUS AR072005 50 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 35 from patent US 5912168.
ACCESSION AR072005
VERSION AR072005.1 GI:7222893
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 50)
AUTHORS Watson,J.D. and Rudert,F.
TITLE CD95 regulatory gene sequences
JOURNAL Patent: US 5912168-A 35 15-JUN-1999;
FEATURES
Location/Qualifiers
source
1. .50
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match
Best Local Similarity 100.0%; Score 28; DB 6; Length 50;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCGGAAGTGCATCCAAATTCAGGTTTC 28
Db 1 GTCGGAAGTGCATCCAAATTCAGGTTTC 28

RESULT 4
AR071999
LOCUS AR071999 25 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 29 from patent US 5912168.
ACCESSION AR071999
VERSION AR071999.1 GI:7222887
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 25)
AUTHORS Watson,J.D. and Rudert,F.
TITLE CD95 regulatory gene sequences
JOURNAL Patent: US 5912168-A 29 15-JUN-1999;
FEATURES
Location/Qualifiers
source
1. .25
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match
Best Local Similarity 100.0%; Score 16; DB 6; Length 25;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCGGAAGTGCATCC 16
Db 10 GTCGGAAGTGCATCC 25

RESULT 5
A62083
LOCUS A62083 36 bp DNA linear PAT 09-MAR-1998
DEFINITION Sequence 18 from Patent WO9714719.
ACCESSION A62083
VERSION A62083.1 GI:3716134
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1
AUTHORS Davis,P.J., Van,D.L., Verhoeijen, Martine,E. and Wilson,S.
TITLE A BIFUNCTIONAL OR BIVALENT ANTIBODY FRAGMENT ANALOGUE
JOURNAL Patent: WO 9714719-A 18 24-APR-1997;
COMMENT UNILEVER NV (NL)
Other publication AU 6873396 970507.
FEATURES
Location/Qualifiers
source
1. .36
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
/clone="PRIMER DBL.5"

ORIGIN
Query Match
Best Local Similarity 79.2%; Score 16; DB 6; Length 36;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 5 GGAAGTGCATCCAAATTCAGGTTTC 28
Db 13 GGAGCTGCATGCAAAATTCATTTC 36

RESULT 6
A66040
LOCUS A66040 36 bp DNA linear PAT 29-MAR-1999
DEFINITION Sequence 22 from Patent WO9738102.
ACCESSION A66040
VERSION A66040.1 GI:4538027
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 36)
AUTHORS Davis,P.J., Van,D.L. and Verhoeijen,M.E.

```

TITLE MULTIVALENT AND MULTISPECIFIC ANTIGEN-BINDING PROTEIN

JOURNAL Patent: WO 9738102-A 22 16-OCT-1997;
UNILEVER PLC (GB)
COMMENT Other publication AU 2507397 19971029.
FEATURES Location/Qualifiers

source

1..36
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
/clone="PRIMER DBL.10"

ORIGIN

Query Match 57.1%; Score 16; DB 6; Length 36;
Best Local Similarity 79.2%; Pred. No. 5.9e+03;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 GGAAGTCATCCAAATTCAGGTTTC 28
|||||
Db 13 GGAGCTGCATGCAAAATTCATTTTC 36

RESULT 7

AR088019 36 bp DNA linear PAT 07-SEP-2000
LOCUS Sequence 18 from patent US 5989830.
DEFINITION

ACCESSION AR088019

VERSION AR088019.1 GI:10014782

KEYWORDS

SOURCE

Unknown.

ORGANISM

Unclassified.

REFERENCE 1 (bases 1 to 36)

AUTHORS Davis,P.James., van der Logt,C.Paul.Erik., Verhoeven,M.Elisa. and

Wilson,S.

TITLE Bifunctional or bivalent antibody fragment analogue

JOURNAL Patent: US 5989830-A 18 23-NOV-1999;

FEATURES Location/Qualifiers

source 1..36

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Query Match 57.1%; Score 16; DB 6; Length 36;
Best Local Similarity 79.2%; Pred. No. 5.9e+03;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 GGAAGTCATCCAAATTCAGGTTTC 28
|||||
Db 13 GGAGCTGCATGCAAAATTCATTTTC 36

RESULT 8

AR154830 36 bp DNA linear PAT 08-AUG-2001
LOCUS Sequence 22 from patent US 6239259.
DEFINITION

ACCESSION AR154830

VERSION AR154830.1 GI:15122883

KEYWORDS

SOURCE

Unknown.

ORGANISM

Unclassified.

REFERENCE 1 (bases 1 to 36)

AUTHORS Davis,P.J., Logt van der,C.P. and Verhoeven,M.E.

TITLE Multivalent and multispecific antigen-binding protein

JOURNAL Patent: US 6239259-A 22 29-MAY-2001;

FEATURES Location/Qualifiers

source 1..36

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Query Match 57.1%; Score 16; DB 6; Length 36;
Best Local Similarity 79.2%; Pred. No. 5.9e+03;

Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 5 GGAAGTCATCCAAATTCAGGTTTC 28
|||||
Db 13 GGAGCTGCATGCAAAATTCATTTTC 36

RESULT 9

AR343590 36 bp DNA linear PAT 17-AUG-2003
LOCUS Sequence 5 from patent US 6579842.
DEFINITION

ACCESSION AR343590

VERSION AR343590.1 GI:33739329

KEYWORDS

SOURCE

Unknown.

ORGANISM

Unclassified.

REFERENCE 1 (bases 1 to 36)

AUTHORS Howell,S., Little,J., Van Der Logt,C.P. and Parry,N.J.

TITLE Method of treating fabrics

JOURNAL Patent: US 6579842-A 5 17-JUN-2003;

FEATURES Location/Qualifiers

source 1..36

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match 57.1%; Score 16; DB 6; Length 36;
Best Local Similarity 79.2%; Pred. No. 5.9e+03;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 GGAAGTCATCCAAATTCAGGTTTC 28
|||||
Db 13 GGAGCTGCATGCAAAATTCATTTTC 36

RESULT 10

AX613512/c 20 bp DNA linear PAT 17-FEB-2003
LOCUS Sequence 4537 from Patent WO2072882.
DEFINITION

ACCESSION AX613512

VERSION AX613512.1 GI:28408941

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

Cullen,P. and Seedorf,U.

AUTHORS Coronary chip

TITLE Patent: WO 02072882-A 4537 19-SEP-2002;

JOURNAL OGHAM GmbH (DE)

FEATURES Location/Qualifiers

source 1..20

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 55.0%; Score 15.4; DB 6; Length 20;
Best Local Similarity 94.1%; Pred. No. 1.1e+04;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 ACTGCATCCAAATTCAG 24
|||||
Db 20 ACTGCTTCCAAATTCAG 4

RESULT 11

AX486149 65 bp DNA linear PAT 16-AUG-2002
LOCUS Sequence 3449 from Patent WO02053728.
DEFINITION

ACCESSION AX486149

```

VERSION      AX486149.1  GI:22320365
KEYWORDS
SOURCE       Candida albicans
ORGANISM     Candida albicans
REFERENCE    1 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
              1 Saccharomycetales; mitosporic Saccharomycetales; Candida.
AUTHORS      Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
TITLE        Gene disruption methodologies for drug target discovery
JOURNAL      Patent: WO 02053728-A 3449 11-JUL-2002;
              Elitra Pharmaceuticals, Inc. (US)
FEATURES     1..65
              Location/Qualifiers
              source
                1..65
                /organism="Candida albicans"
                /mol_type="unassigned DNA"
                /db_xref="taxon:5476"
ORIGIN
Query Match      54.3%; Score 15.2; DB 6; Length 65;
Best Local Similarity 85.0%; Pred. No. 1.7e+04;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 4 TGGAACTGCATCCAAATTC A 23
    ||||| ||||| ||||| |||||
DB 18 TGGACCAGCATCCAAATTC A 37
    ||||| ||||| ||||| |||||

RESULT 12
AX386437/c
LOCUS          AX386437      82 bp      DNA      linear      PAT 19-MAR-2002
DEFINITION     Sequence 1365 from Patent WO0214500.
ACCESSION      AX386437
VERSION        AX386437.1  GI:19579567
KEYWORDS       Homo sapiens (human)
SOURCE         Homo sapiens
ORGANISM       Homo sapiens
REFERENCE      1 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
              1 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS        Escobedo,J., Garcia,P.D., Sudduth-Klinger,J., Reinhard,C.,
              1 Radazzo,F., Lamson,G., Scott,E.M., Zhang,G., Kassam,A., Pot,D. and
              1 Labat,I.
TITLE          Human genes and gene expression products
JOURNAL        Patent: WO 0214500-A 1365 21-FEB-2002;
              CHIRON CORPORATION (US); Hyseq Inc. (US)
FEATURES       1..82
              Location/Qualifiers
              source
                1..82
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"
ORIGIN
Query Match      54.3%; Score 15.2; DB 6; Length 82;
Best Local Similarity 85.0%; Pred. No. 1.7e+04;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 7 AACTGCATCCAAATTCAGGT 26
    ||||| ||||| ||||| |||||
DB 26 ACCTGCTCTAAATTCAGGT 7
    ||||| ||||| ||||| |||||

RESULT 13
AF357525
LOCUS          AF357525      84 bp      mRNA      linear      PRI 11-JUN-2001
DEFINITION     Homo sapiens Bcl-XL-binding protein c21 mRNA, partial cds.
ACCESSION      AF357525
VERSION        AF357525.1  GI:14335461
KEYWORDS       Homo sapiens (human)
SOURCE         Homo sapiens
ORGANISM       Homo sapiens
REFERENCE      1 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
              1 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              1 (bases 1 to 84)

```

```

AUTHORS      Hammond,P.W., Alpin,J., Rise,C.E., Wright,M. and Kreider,B.L.
TITLE        In vitro selection and characterization of Bcl-X(L)-binding
              proteins from a mix of tissue-specific mRNA display libraries
JOURNAL      J. Biol. Chem. 276 (24), 20898-20906 (2001)
MEDLINE      21293069
PUBMED       11283018
REFERENCE    2 (bases 1 to 84)
AUTHORS      Hammond,P.W., Alpin,J., Rise,C.E., Wright,M. and Kreider,B.L.
TITLE        Direct Submission
JOURNAL      Submitted (08-MAR-2001) Molecular Biology, Phyllos, Inc., 128 Spring
              St., Lexington, MA 02421, USA
FEATURES     1..84
              Location/Qualifiers
              source
                1..84
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /tissue_type="kidney"
                <1..>84
                /note="EXLBC21"
                /codon_start=1
                /product="Bcl-XL-binding protein c21"
                /protein_id="AAK60629.1"
                /db_xref="GI:14335462"
                /translation="IAQATKATIDKWNCKIKLIFYTSKEAS"
ORIGIN
Query Match      54.3%; Score 15.2; DB 9; Length 84;
Best Local Similarity 85.0%; Pred. No. 1.7e+04;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 4 TGGAACTGCATCCAAATTC A 23
    ||||| ||||| ||||| |||||
DB 34 TGGAACTGCATCCAAATTC A 53
    ||||| ||||| ||||| |||||

RESULT 14
AX769938/c
LOCUS          AX769938      99 bp      DNA      linear      PAT 03-JUL-2003
DEFINITION     Sequence 3 from Patent WO03016341.
ACCESSION      AX769938
VERSION        AX769938.1  GI:32437573
KEYWORDS       Homo sapiens (human)
SOURCE         Homo sapiens
ORGANISM       Homo sapiens
REFERENCE      1 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
              1 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              1 Hubert,N. and Hentze,M.
              1 Regulation of iron uptake
              1 Patent: WO 03016341-A 3 27-FEB-2003;
              1 European Molecular Biology Laboratory (DE)
              1 Location/Qualifiers
              source
                1..99
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"
ORIGIN
Query Match      54.3%; Score 15.2; DB 6; Length 99;
Best Local Similarity 71.4%; Pred. No. 1.7e+04;
Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 GTCTGGAATGCATCCAAATTCAGGTTTC 28
    ||||| ||||| ||||| |||||
DB 89 GTGTGGAGCTGCCTCCGCTCTTCAGTGC 62
    ||||| ||||| ||||| |||||

RESULT 15
AX769969/c
LOCUS          AX769969      99 bp      DNA      linear      PAT 02-JUL-2003
DEFINITION     Sequence 34 from Patent WO03016341.
ACCESSION      AX769969
VERSION        AX769969.1  GI:32437600
KEYWORDS
SOURCE
ORGANISM
REFERENCE

```


KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Hubert, N. and Hentze, M.
 TITLE Regulation of iron uptake
 JOURNAL Patent: WO 03016341-A 34 27-FEB-2003;
 European Molecular Biology Laboratory (DE)
 FEATURES Location/Qualifiers
 source
 1..99
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Exon 1A of promoter region of human DMT1"
 ORIGIN
 Query Match 54.3%; Score 15.2; DB 6; Length 99;
 Best Local Similarity 71.4%; Pred. No. 1.7e+04;
 Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1 GTCTGGAACTGCATCCAAATTCAGGTTTC 28
 ||||| ||||| ||||| ||||| |||||
 Db 89 GTGTGGAGCTGCTCCGCTTCAGCTGC 62

Search completed: September 1, 2004, 23:06:52
 Job time : 1398 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2004, 21:27:22 ; Search time 231 Seconds
(without alignments)
514.933 Million cell updates/sec

Title: US-10-028-415-2
Perfect score: 28
Sequence: 1 gcttgaactgcattcaaaattcagggttc 28

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 3485926

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001as:*
- 5: Geneseq2001bs:*
- 6: Geneseq2002s:*
- 7: Geneseq2003as:*
- 8: Geneseq2003bs:*
- 9: Geneseq2003cs:*
- 10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28	100.0	28	2 AAV20897	AAV20897 CD95 regu
2	28	100.0	28	2 AAV23143	AAV23143 CD95 regu
3	28	100.0	28	6 ABK88709	ABK88709 Human CD9
4	28	100.0	28	6 ABK88700	ABK88700 Transcrip
5	28	100.0	50	2 AAV23172	AAV23172 CD95 enha
6	17.2	61.4	65	6 ABN55985	ABN55985 Mouse spl
7	16.6	59.3	60	6 ABN37069	ABN37069 Human spl
8	16.4	58.6	100	7 ACD75131	ACD75131 E. coli K
9	16	57.1	25	2 AAV23161	AAV23161 Primer -1
10	16	57.1	36	2 AAT66152	AAT66152 Primer DB
11	16	57.1	36	2 AAT94605	AAT94605 Construct
12	16	57.1	36	7 AAL53506	AAL53506 Serotonin
13	15.8	56.4	50	6 ABZ04805	ABZ04805 Human leu
14	15.6	55.7	60	6 ABN39933	ABN39933 Human spl
15	15.6	55.7	65	6 ABN28998	ABN28998 Rat splic
16	15.6	55.7	100	7 ACD75130	ACD75130 E. coli K
17	15.4	55.0	60	6 ABN34583	ABN34583 Human spl
18	15.2	54.3	40	2 AAT14497	AAT14497 PelB secr
19	15.2	54.3	65	6 ABZ29366	ABZ29366 Candida g
20	15.2	54.3	66	2 AAT21089	AAT21089 Human gen
21	15.2	54.3	82	6 ABN61398	ABN61398 Human can
22	15.2	54.3	84	7 ABT14837	ABT14837 Human Bcl
23	15.2	54.3	96	2 AAT47779	AAT47779 Polyclona

C 24	15.2	54.3	99	7 ABQ80179	ABQ80179 Human DMT
C 25	15.2	54.3	100	7 ACD69282	ACD69282 E. coli K
C 26	15	53.6	24	3 AAC67600	AAC67600 Alzheimer
C 27	15	53.6	24	3 ADD43523	ADD43523 Human mit
C 28	15	53.6	28	3 AAA15192	AAA15192 PCR prime
C 29	15	53.6	33	6 ABL41320	ABL41320 Human GTP
C 30	15	53.6	43	2 AAQ62546	AAQ62546 Probe spe
C 31	15	53.6	68	2 AAQ71837	AAQ71837 pBROC531
C 32	15	53.6	68	2 AAQ71836	AAQ71836 pBROC531
C 33	15	53.6	68	2 AAT18305	AAT18305 Nucleotid
C 34	15	53.6	68	2 AAT18305	AAT18305 Nucleotid
C 35	15	53.6	83	2 AAV73250	AAV73250 C. utilis
C 36	14.8	52.9	27	2 AAZ30800	AAZ30800 SCBV viru
C 37	14.8	52.9	60	6 ABN46107	ABN46107 Human spl
C 38	14.8	52.9	60	6 ABN47820	ABN47820 Human spl
C 39	14.8	52.9	60	6 ABN46731	ABN46731 Human spl
C 40	14.8	52.9	90	6 ABN69180	ABN69180 Streptoco
C 41	14.8	52.9	100	7 ACD80222	ACD80222 E. coli K
C 42	14.6	52.1	41	6 ABZ48776	ABZ48776 Human ald
C 43	14.6	52.1	41	6 ABZ46265	ABZ46265 Human ald
C 44	14.6	52.1	50	4 AAL31606	AAL31606 Human SNP
C 45	14.6	52.1	50	4 AAL31607	AAL31607 Human SNP

ALIGNMENTS

RESULT 1

AAV20897

ID AAV20897 standard; DNA; 28 BP.

AC AAV20897;

DT 23-JUL-1998 (first entry)

DE CD95 regulatory region.

KW CD95; regulatory region; silencer; enhancer; apoptosis; inhibitor;
cancer; viral infection; neurodegeneration; autoimmune disease;
Gene therapy; transcription factor; ss.

OS Homo sapiens.

PN WO9808965-A2.

PD 05-MAR-1998.

PF 29-AUG-1997; 97WO-NZ000107.

PR 30-AUG-1996; 96US-00713557.

PP (GENE-) GENESIS RES & DEV CORP LTD.

PI Watson JD, Rudert F;

PS WPI; 1998-179445/16.

XX New regulatory regions from the CD95 gene and transcription factors that
interact with them - for control of apoptosis, e.g. in treatment of
cancer, viral infection, neurodegeneration and autoimmune disease.
Claim 1; Page 36; 60pp; English.

CC This sequence represents a regulatory region from the CD95 gene, and is a
nucleic acid of the invention. The regulatory regions (silencers or
enhancers) are involved in apoptosis, i.e. inhibition of CD95 expression
will inhibit apoptosis. Regulation of apoptosis is useful in treatment of
cancer, (retroviral infection, neurodegeneration and autoimmune disease,
e.g. by gene therapy for expressing transcription factors or expression
of antisense sequences to inhibit transcription factor production. The
regulatory nucleic acids and transcription factors are also useful for
studying regulation of CD95 in vitro or in vivo, to screen for modulators
and as probes to isolate related genes

```

XX SQ Sequence 28 BP; 7 A; 7 C; 6 G; 8 T; 0 U; 0 Other;
Query Match 100.0%; Score 28; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.0064;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGTGAAGTGCATCCAAATTCAGGTTTC 28
Db 1 GTCGTGAAGTGCATCCAAATTCAGGTTTC 28

RESULT 2
AAV23143/c
ID AAV23143 standard; DNA; 28 BP.
AC AAV23143;
XX
XX DT 23-JUL-1998 (first entry)
XX DE CD95 regulatory region.
XX KW CD95; regulatory region; silencer; enhancer; apoptosis; inhibitor;
XX KW cancer; viral infection; neurodegeneration; autoimmune disease;
XX KW gene therapy; transcription factor; ss.
XX OS Homo sapiens.
XX PN WO9808965-A2.
XX PD 05-MAR-1998.
XX FF 29-AUG-1997; 97WO-NZ000107.
XX PR 30-AUG-1996; 96US-00713557.
XX PA (GENE-) GENESIS RES & DEV CORP LTD.
XX PI Watson JD, Rudert F;
XX DR WPI; 1998-179445/16.
XX PT New regulatory regions from the CD95 gene and transcription factors that
XX PT interact with them - for control of apoptosis, e.g. in treatment of
XX PT cancer, viral infection, neurodegeneration and autoimmune disease.
XX PS Claim 1; Page 38; 60pp; English.
XX CC This sequence represents a regulatory region from the CD95 gene, and is a
XX CC nucleic acid of the invention. The regulatory regions (silencers or
XX CC enhancers) are involved in apoptosis, i.e. inhibition of CD95 expression
XX CC will inhibit apoptosis. Regulation of apoptosis is useful in treatment of
XX CC cancer, (retroviral infection, neurodegeneration and autoimmune disease,
XX CC e.g. by gene therapy for expressing transcription factors or expression
XX CC of antisense sequences to inhibit transcription factor production. The
XX CC regulatory nucleic acids and transcription factors are also useful for
XX CC studying regulation of CD95 in vitro or in vivo, to screen for modulators
XX CC and as probes to isolate related genes
XX SQ Sequence 28 BP; 8 A; 6 C; 7 G; 7 T; 0 U; 0 Other;
Query Match 100.0%; Score 28; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.0064;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGTGAAGTGCATCCAAATTCAGGTTTC 28
Db 28 GTCGTGAAGTGCATCCAAATTCAGGTTTC 1

RESULT 3
ABK88709/c
ID ABK88709 standard; DNA; 28 BP.

```

```

XX AC ABK88709;
XX DT 07-OCT-2002 (first entry)
XX DE Human CD95 gene transcription silencer probe #1.
XX KW Human; apoptotic cell death; proteinaceous transcription factor;
XX KW regulation of gene transcription; apoptosis; p53; CD95; TRA;
XX KW transcriptional regulator of apoptosis; Y-box family; YB-1; cancer;
XX KW tumour cell; embryonic cell; nervous system; intracellular pathogen;
XX KW DNA-damaging agent; retroviral infection; neurodegenerative disorder;
XX KW immune system dysfunction; anti-tumour; cytostatic; hCD95;
XX KW transcription silencer region; probe; ss.
XX OS Homo sapiens.
XX PN WO200244363-A1.
XX PD 06-JUN-2002.
XX PF 28-NOV-2001; 2001WO-NZ000287.
XX PR 28-NOV-2000; 2000US-00724809.
XX PA (GENE-) GENESIS RES & DEV CORP LTD.
XX PI Lasham A, Watson JD;
XX DR WPI; 2002-557540/59.
XX PT Modulating p53-mediated apoptotic cell death in a population of cells, by
XX PT modulating the amount of a transcriptional regulator of apoptosis
XX PT available to bind to a target polynucleotide in the cells.
XX PS Example 1; Page 54; 62pp; English.
XX CC The present invention relates to methods for modulating apoptotic cell
XX CC death using proteinaceous transcription factors that regulate the
XX CC transcription of genes encoding proteins involved in apoptosis (e.g. CD95
XX CC and p53). The methods involve modulating the amount of a transcriptional
XX CC regulator of apoptosis (TRA) available to bind to a target polynucleotide
XX CC in the cells, where TRA is a member of the Y-box nucleic acid binding
XX CC family of polypeptides (e.g. YB-1). The methods of the invention are
XX CC useful for modulating apoptotic cell death in a population of cells,
XX CC where the cells are selected from tumour cells, cells of the immune
XX CC system, embryonic cells, cells of the nervous system, or cells infected
XX CC with intracellular pathogens. The methods are also useful for increasing
XX CC the sensitivity of tumour cells to a DNA-damaging agent, and for
XX CC increasing sensitivity to apoptosis in a population of cells harbouring
XX CC intracellular pathogens. The methods are useful for screening an
XX CC apoptosis modulatory agent that modulates the binding of TRA. The methods
XX CC for regulating apoptosis can be used therapeutically and prophylactically
XX CC for various disorders such as cancer, viral and retroviral infections,
XX CC neurodegenerative disorders, and immune system dysfunction. The present
XX CC sequence represents a human CD95 (hCD95) gene transcription silencer
XX CC probe
XX SQ Sequence 28 BP; 8 A; 6 C; 7 G; 7 T; 0 U; 0 Other;
Query Match 100.0%; Score 28; DB 6; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.0064;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGTGAAGTGCATCCAAATTCAGGTTTC 28
Db 28 GTCGTGAAGTGCATCCAAATTCAGGTTTC 1

RESULT 4
ABK88700
ID ABK88700 standard; DNA; 28 BP.

```

```
AC ABK8700;
XX
XX 07-OCT-2002 (first entry)
XX
XX Transcription silencer region (S1) of human CD95 gene.
XX
XX Human; apoptotic cell death; proteinaceous transcription factor;
XX regulation of gene transcription; apoptosis; p53; CD95; TRA;
XX transcriptional regulator of apoptosis; I-box family; YB-1; cancer;
XX tumour cell; embryonic cell; nervous system; intracellular pathogen;
XX DNA-damaging agent; retroviral infection; neurodegenerative disorder;
XX immune system dysfunction; anti-tumour; cytostatic; S1; hCD95;
XX transcription silencer region; ds.
XX
XX Homo sapiens.
XX
XX WO200244363-A1.
XX
XX 06-JUN-2002.
XX
XX 28-NOV-2001; 2001WO-NZ000287.
XX
XX 28-NOV-2000; 2000US-00724809.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX
XX Lasham A, Watson JD;
XX
XX WPI; 2002-557540/59.
XX
XX Modulating p53-mediated apoptotic cell death in a population of cells, by
XX modulating the amount of a transcriptional regulator of apoptosis
XX available to bind to a target polynucleotide in the cells.
XX
XX Example 1; Page 53; 62pp; English.
XX
XX The present invention relates to methods for modulating apoptotic cell
XX death using proteinaceous transcription factors that regulate the
XX transcription of genes encoding proteins involved in apoptosis (e.g. CD95
XX and p53). The methods involve modulating the amount of a transcriptional
XX regulator of apoptosis (TRA) available to bind to a target polynucleotide
XX in the cells, where TRA is a member of the Y-box nucleic acid binding
XX family of polypeptides (e.g. YB-1). The methods of the invention are
XX useful for modulating apoptotic cell death in a population of cells,
XX where the cells are selected from tumour cells, cells of the immune
XX system, embryonic cells, cells of the nervous system, or cells infected
XX with intracellular pathogens. The methods are also useful for increasing
XX the sensitivity of tumour cells to a DNA-damaging agent, and for
XX increasing sensitivity to apoptosis in a population of cells harbouring
XX intracellular pathogens. The methods are useful for screening an
XX apoptosis modulatory agent that modulates the binding of TRA. The methods
XX for regulating apoptosis can be used therapeutically and prophylactically
XX for various disorders such as cancer, viral and retroviral infections,
XX neurodegenerative disorders, and immune system dysfunction. The present
XX sequence represents the transcription silencer region (S1) of the human
XX CD95 (hCD95) gene.
XX
XX Sequence 28 BP; 7 A; 7 C; 6 G; 8 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 28; DB 6; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.0064;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTCTGGAACCTGCATCCAAATTCAGGTTTC 28
Db 1 GTCTGGAACCTGCATCCAAATTCAGGTTTC 28
RESULT 5
AAV23172
ID AAV23172 standard; DNA; 50 BP.
XX
XX AAV23172;
AC
```

```
XX 23-JUL-1998 (first entry)
XX
XX CD95 enhancer binding site.
XX
XX CD95; regulatory region; silencer; enhancer; apoptosis; inhibitor;
XX cancer; viral infection; neurodegeneration; autoimmune disease;
XX gene therapy; transcription factor; ss.
XX
XX Homo sapiens.
XX
XX WO9808965-A2.
XX
XX 05-MAR-1998.
XX
XX 29-AUG-1997; 97WO-NZ000107.
XX
XX 30-AUG-1996; 96US-00713557.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX
XX Watson JD, Rudert F;
XX
XX WPI; 1998-179445/16.
XX
XX New regulatory regions from the CD95 gene and transcription factors that
XX interact with them - for control of apoptosis, e.g. in treatment of
XX cancer, viral infection, neurodegeneration and autoimmune disease.
XX
XX Disclosure; Page 43; 60pp; English.
XX
XX This sequence represents a regulatory region from the CD95 gene, and is a
XX nucleic acid of the invention. The regulatory regions (silencers or
XX enhancers) are involved in apoptosis, i.e. inhibition of CD95 expression
XX will inhibit apoptosis. Regulation of apoptosis is useful in treatment of
XX cancer, (retro)viral infection, neurodegeneration and autoimmune disease,
XX e.g. by gene therapy for expressing transcription factors or expression
XX of antisense sequences to inhibit transcription factor production. The
XX regulatory nucleic acids and transcription factors are also useful for
XX studying regulation of CD95 in vitro or in vivo, to screen for modulators
XX and as probes to isolate related genes.
XX
XX Sequence 50 BP; 16 A; 10 C; 10 G; 14 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 28; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 0.007;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTCTGGAACCTGCATCCAAATTCAGGTTTC 28
Db 1 GTCTGGAACCTGCATCCAAATTCAGGTTTC 28
RESULT 6
ABN55985/C
ID ABN55985 standard; DNA; 65 BP.
XX
XX ABN55985;
AC
XX
XX 15-JUL-2002 (first entry)
XX
XX Mouse spliced transcript detection oligonucleotide SEQ ID NO:28733.
XX
XX Human; mouse; rat; splice transcript; detection; RNA transcript;
XX splice variant; transcriptome; oligonucleotide library; ss.
XX
XX Mus musculus.
XX
XX WO200210449-A2.
XX
XX 07-FEB-2002.
XX
XX 20-JUL-2001; 2001WO-IB001903.
XX
XX
```

[illegible]

PA (MWGB-) MWG-BIOTECH AG.
 XX
 PI Donner H, Drescher B, Huber A, Weber J;
 XX WPI; 2003-241155/24.
 XX
 DR Biochip containing probes complementary with open reading frames in
 PT Escherichia coli K12, useful for detecting gene expression and expression
 PT patterns.
 XX
 PS Claim 3; Page 1002; 2004pp; German.
 XX
 CC This invention describes a novel biochip comprising probe spots, each
 CC containing many identical probes. The probes are nucleotide sequences of
 CC 30-80 bases, are prepared ex situ from synthetic oligonucleotides and at
 CC least one includes a segment of at least 20 bases identical with, or
 CC complementary to, a segment of an open reading frame (orf) of Escherichia
 CC coli K12. The biochip is used for specific detection of gene expression
 CC in K12 and for determining the gene expression pattern, e.g. for
 CC diagnostic determination of which E. coli strains are present in the gut,
 CC and to determine the effects of e.g. growth media on gene expression. The
 CC biochip provides as comprehensive as possible detection of the K12
 CC genome, with simultaneous analysis of many different genes with a single
 CC device, and comparison of gene expression between K12 and its mutants or
 CC other E. coli strains in a single experiment. Apart from qualitative and
 CC quantitative information about gene expression, it also allows
 CC measurements of population densities for the various strains. The use of
 CC synthetic oligonucleotides for preparation of probes allows free
 CC variation in probe length and ensures high purity (and thus selectivity,
 CC reactivity and reproducibility); also synthetic probes are generally
 CC shorter than probes prepared by polymerase chain reaction. ACD68731 to
 CC ACD81540 represent oligonucleotide probes used with the biochip described
 CC in the invention
 XX
 SQ Sequence 100 BP; 19 A; 20 C; 31 G; 30 T; 0 U; 0 Other;
 Query Match 58.6%; Score 16.4; DB 7; Length 100;
 Best Local Similarity 76.9%; Pred. No. 1.4e+03;
 Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 3 CTGGAACTGCATCCCAATTCAGGTC 28
 DB 88 CCGGAACGGCATCCCACTTCAGAAATC 63
 RESULT 9
 AAV23161
 ID AAV23161 standard; DNA; 25 BP.
 AC AAV23161;
 DT 23-JUL-1998 (first entry)
 DE Primer -1035/-1/CAT for human CD95 gene.
 XX
 KW CD95; regulatory region; silencer; enhancer; apoptosis; inhibitor;
 KW cancer; viral infection; neurodegeneration; autoimmune disease; human;
 KW gene therapy; transcription factor; PCR primer; ss.
 XX
 OS Synthetic.
 OS Homo sapiens.
 PN WO9808965-A2.
 XX
 PD 05-MAR-1998.
 XX
 PF 29-AUG-1997; 97WO-NZ000107.
 XX
 PR 30-AUG-1996; 96US-00713557.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 XX
 PI Watson JD, Rudert F;
 WPI; 1998-179445/16.
 New regulatory regions from the CD95 gene and transcription factors that
 interact with them - for control of apoptosis, e.g. in treatment of
 cancer, viral infection, neurodegeneration and autoimmune disease.
 Disclosure; Page 42; 60pp; English.
 This sequence is a primer for the CD95 gene, and is used to isolate the
 regulatory nucleic acids of the invention. The regulatory regions
 (silencers or enhancers) are involved in apoptosis, e.g. inhibition of
 CD95 expression will inhibit apoptosis. Regulation of apoptosis is useful
 in treatment of cancer, (retro)viral infection, neurodegeneration and
 autoimmune disease, e.g. by gene therapy for expressing transcription
 factors or expression of antisense sequences to inhibit transcription
 factor production. The regulatory nucleic acids and transcription factors
 are also useful for studying regulation of CD95 in vitro or in vivo, to
 screen for modulators and as probes to isolate related genes
 Sequence 25 BP; 5 A; 7 C; 7 G; 6 T; 0 U; 0 Other;
 Query Match 57.1%; Score 16; DB 2; Length 25;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTCTGGAACTGCATCC 16
 DB 10 GTCTGGAACTGCATCC 25
 RESULT 10
 AAT66152
 ID AAT66152 standard; DNA; 36 BP.
 AC AAT66152;
 DT 14-JUL-1997 (first entry)
 DE Primer DBL.5 used in vector pGOSA.E construction.
 XX
 KW Bispecific antibody; bivalent antibody; double head;
 KW antibody engineering; glucose oxidase; Streptococcus sanguis; GOSA.E;
 KW vector; polymerase chain reaction; PCR; primer; ss.
 XX
 OS Synthetic.
 XX WO9714719-A1.
 XX
 PD 24-APR-1997.
 XX
 PF 14-AUG-1996; 96WO-EP003605.
 XX
 PR 16-OCT-1995; 95EP-00307332.
 XX
 PA (UNIL) UNILEVER NV.
 PA (UNIL) UNILEVER PLC.
 XX
 PI Davis PJ, Van Der Logt CPE, Verhoeijen ME, Wilson S;
 WPI; 1997-245049/22.
 Bi-specific or bivalent antibody fragment analogues comprise complex of
 two polypeptide chains containing respectively two heavy and two light
 chain variable domains - useful therapeutically, in immunoassays, for
 purification etc.
 Example 1; Page 18; 121pp; English.
 PCR primers (AAT66148-63) were utilised in the construction of 'double
 head' expression vector pGOSA.E (see also AAT66147), which encodes a
 bispecific antibody (see also AAW09818-20) for glucose oxidase and
 Streptococcus sanguis. The primers were used to generate restriction

CC sites to facilitate the cloning of DNA fragments utilised in vector
 CC construction. Primer DBL.5 (AA766152) contg. an NheI site was designed to
 CC match the sequence upstream of the ribosome binding site of VL.4715 in
 CC Fv4715 and to introduce an NheI site
 XX
 SQ Sequence 36 BP; 11 A; 7 C; 7 G; 11 T; 0 U; 0 Other;

Query Match 57.1%; Score 16; DB 2; Length 36;
 Best Local Similarity 79.2%; Pred. No. 1.9e+03;
 Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 GGAAGTCATCCAAATTCAGGTTTC 28
 ||| ||||| ||||| ||||| |||||
 Db 13 GGAGCTGCATGCAAAATTCATTTTC 36

RESULT 11

AA794605
 ID AAT94605 standard; DNA; 36 BP.

XX AC AAT94605;

XX DT 24-APR-1998 (first entry)

DE Construction of trispecific pSV constructs using primer DBL.10.

XX Multivalent antigen binding protein; variable domain; binding site;
 KW diagnosis; therapy; medicine; immunoassay; target; trispecific;
 KW pSV construct; PCR primer; ss.
 XX OS Synthetic.

XX WO9738102-A1.
 XX 16-OCT-1997.

XX 26-MAR-1997; 97WO-EF001609.
 XX 04-APR-1996; 96EP-00302412.

XX (UNIL) UNILEVER PLC.
 XX (UNIL) UNILEVER NV.

PI Davis PJ, Van Der Logt CPE, Verhoeven ME;
 XX WPI; 1997-512719/47.

XX Multivalent antigen binding protein - comprises antibody variable
 PT domains, linked by their association, where each associated variable
 PT domain pair forms an antigen binding site.

XX Disclosure; Page 41; 100pp; English.

XX This PCR primer is used in the construction of the trispecific pSV
 CC constructs which are used to construct a novel multivalent antigen
 CC binding protein. This novel protein comprises 2 polypeptides, each
 CC comprising, in series, 3 or more variable domains of an antibody heavy or
 CC light chain respectively, which are linked by association of the
 CC respective heavy and light chain variable domains, where each associated
 CC variable domain pair forms an antigen binding site. The multivalent
 CC antigen binding protein can be used in medicine, diagnosis, therapy, an
 CC immunoassay method or for purification. It may be used in targeting a
 CC tumour cell with natural killer cells and cytotoxic agents. It may be
 CC used to target cell killing enzymes, e.g. oxidases and peroxidases, to a
 CC species with which is an antigenic component of a dental plaque, such as
 CC S. sanguis or S. mutans. The antigen binding protein may conveniently be
 CC purified straight from the supernatant using conventional purification
 CC techniques. As the protein is self-assembling, there is no need to purify
 CC individual subunits prior to coupling as in existing techniques

SQ .Sequence 36 BP; 11 A; 7 C; 7 G; 11 T; 0 U; 0 Other;

Query Match 57.1%; Score 16; DB 2; Length 36;

Best Local Similarity 79.2%; Pred. No. 1.9e+03;
 Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 GGAAGTCATCCAAATTCAGGTTTC 28
 ||| ||||| ||||| ||||| |||||
 Db 13 GGAGCTGCATGCAAAATTCATTTTC 36

RESULT 12

AA753506
 ID AAL53506 standard; DNA; 36 BP.

XX AC AAL53506;

XX DT 30-JAN-2003 (first entry)

XX Serotonin receptor internal oligo probe SEQ ID No 10.

XX Immunomodulator; antirheumatic; antiarthritic; immunosuppressive;
 KW haemostatic; antiinflammatory; antiulcer; neuroprotective; antithyroid;
 KW antidiabetic; dermatological; antipsoriatic; gynaecological; vasotropic;
 KW anti-HIV; immune response; inhibitor; serotonin; serotonin receptor;
 KW CD-4; CD-8; T cell; B cell; autoimmune disease; fulminant AIDS;
 KW serotonin receptor oligonucleotide probe; ss.

XX OS Unidentified.

XX WO200278643-A2.

XX 10-OCT-2002.

XX 29-MAR-2002; 2002WO-US009993.

XX 30-MAR-2001; 2001US-0280296P.

XX 25-OCT-2001; 2001US-0345295P.

XX 31-JAN-2002; 2002US-0353883P.

XX (PHIL-) PHILADELPHIA HEALTH & EDUCATION CORP.

XX Jameson BA, Tretiakova AS, Albert R, Davidson HC;

XX WPI; 2003-040619/03.

XX Modulating immune response in mammal in treatment of e.g. multiple
 PT sclerosis, myasthenia gravis, chronic neutropenia, Crohn's disease,
 PT endometriosis, involves administering inhibitor of interaction of
 PT serotonin with serotonin receptor.

XX Example 2; Page 21; 172pp; English.

XX The invention relates to a discovery that modulating an immune response
 CC in a mammal involves administering an inhibitor of the interaction of
 CC serotonin with a serotonin receptor. The invention is useful for
 CC modulating (e.g. inhibiting) an immune response (such as CD-4 or CD-8
 CC dependent immune response); for inhibiting an immune reaction or response
 CC mediated by activation of serotonin receptor on an immune cell (such as T
 CC cell and B cell) due to the activation of the serotonin receptor on the
 CC cell; for modulating an immune response of an autoimmune disease (such as
 CC myasthenia gravis, idiopathic inflammatory myopathy, chronic neutropenia,
 CC rheumatoid arthritis, idiopathic thrombocytopenia purpura, autoimmune
 CC haematologic syndromes, antiphospholipid antibody syndromes, inflammatory
 CC bowel disease, Crohn's disease, ulcerative colitis, myocarditis, Guillain
 CC -Barre syndrome, vasculitis, multiple sclerosis, neuromyelitis optica
 CC (Devic's syndrome), lymphocytic hypophysitis, Grave's disease, Addison's
 CC disease, hypoparathyroidism, type I diabetes, systemic lupus erythematosus,
 CC pemphigus vulgaris, bullous pemphigoid, psoriasis, psoriatic arthritis,
 CC endometriosis, autoimmune orchitis, autoimmune erectile dysfunction,
 CC sarcoidosis, Wegener's granulomatosis, autoimmune deafness, Sjogren's
 CC disease, autoimmune uveoretinitis, interstitial cystitis, Goodpasture's
 CC syndrome, and fibromyalgia); for inhibiting a secondary immune response,
 CC in a mammal (preferably a human); and for inducing apoptosis or death in
 CC a cell or affecting a cell cycle process in a cell expressing a serotonin
 CC receptor by inhibiting transmission of a serotonin signal via a serotonin

CC receptor. The invention is also useful for treating fulminant AIDS. This
 CC polynucleotide sequence represents a serotonin receptor oligonucleotide
 CC probe relating to the invention

XX Sequence 36 BP; 12 A; 13 C; 3 G; 8 T; 0 U; 0 Other;

Query Match 57.1%; Score 16; DB 7; Length 36;
 Best Local Similarity 79.2%; Pred. No. 1.9e+03;
 Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 TCTGGAACCTGCATCCAAATTCAGG 25
 |||||
 Db 10 TCAGCAACTACCTCCAAACTCAGG 33

RESULT 13

ABZ04805
 ID ABZ04805 standard; DNA; 50 BP.

XX AC ABZ04805;

XX 09-JAN-2003 (first entry)

XX Human leukocyte gene expression profiling probe SEQ ID NO 4796.

XX T7; leukocyte; gene expression profiling; allograft rejection;
 KW atherosclerosis; congestive heart failure; systemic lupus erythematosus;
 KW rheumatoid arthritis; osteoarthritis; cytomegalovirus; infection; probe;
 KW ss.

XX Homo sapiens.

XX WO200257414-A2.

XX 25-JUL-2002.

XX 22-OCT-2001; 2001WO-US047856.

XX 20-OCT-2000; 2000US-0241994P.

XX 08-JUN-2001; 2001US-0296764P.

XX (BIOC-) BIOCARDIA INC.

XX Wohlgemuth J, Fry X, Marcuk G, Altman P, Prentice J, Phillips J;

XX Ly N, Woodward R, Quertemous T, Johnson F;

XX WPI; 2002-636525/68.

XX New system for leukocyte expression profiling, diagnosing a disease, or
 PT monitoring (the rate of) progression of a disease, e.g. atherosclerosis
 PT or congestive heart failure, comprises diagnostic oligonucleotides.

XX Claim 1; Page 481; Opp; English.

XX The invention relates to a system for detecting gene expression, which
 CC comprises one or two isolated DNA molecules that detect expression of a
 CC gene, where the gene corresponds to any of 8143 oligonucleotides
 CC (ABZ0010-ABZ08152) each having 50 base pairs (bp). The system is useful
 CC for leukocyte expression profiling. It is particularly useful for
 CC diagnosing a disease, monitoring (rate of) progression of a disease,
 CC predicting therapeutic outcome, determining prognosis for a patient,
 CC predicting disease complications in an individual or monitoring response
 CC to treatment in an individual. The diseases include cardiac allograft
 CC rejection, kidney allograft rejection, liver allograft rejection,
 CC atherosclerosis, congestive heart failure, systemic lupus erythematosus,
 CC rheumatoid arthritis, osteoarthritis or cytomegalovirus infection

XX Sequence 50 BP; 15 A; 13 C; 7 G; 15 T; 0 U; 0 Other;

Query Match 56.4%; Score 15.8; DB 6; Length 50;

Best Local Similarity 74.1%; Pred. No. 2.4e+03;

Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 TCTGGAACCTGCATCCAAATTCAGGTTTC 28
 |||||
 Db 24 TCTGGTACTAGTACAAATTCGGTTTC 50

RESULT 14

ABN39933/C

ID ABN39933 standard; DNA; 60 BP.

XX AC ABN39933;

XX 15-JUL-2002 (first entry)

XX Human spliced transcript detection oligonucleotide SEQ ID NO:12681.

XX Human; mouse; rat; splice transcript; detection; RNA transcript;

XX splice variant; transcriptome; oligonucleotide library; ss.

XX Homo sapiens.

XX WO200210449-A2.

XX 07-FEB-2002.

XX 20-JUL-2001; 2001WO-IB001903.

XX 28-JUL-2000; 2000US-0221607P.

XX 02-MAY-2001; 2001US-0287724P.

XX (COMP-) COMPUGEN INC.

XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;

XX WPI; 2002-257383/30.

XX New oligonucleotide libraries comprising oligonucleotides which
 PT selectively hybridize to mRNAs transcribed from a transcription unit of a
 PT genome, useful for detecting tissue-, pathology-, and developmental-
 PT specific genes.
 PS Example 1; SEQ ID NO 12681; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting
 CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-
 CC)transcriptome comprises messenger RNAs transcribed from multiple
 CC transcription units that populate a genome. The library comprises several
 CC oligonucleotides, each capable of hybridising selectively to a set of
 CC messenger RNAs transcribed from a given transcription unit of the genome,
 CC which encodes one or more messenger RNA splice variants. The
 CC oligonucleotide libraries are useful for detecting mRNAs from a
 CC biological sample, in expression profiling studies, in qualitatively or
 CC quantitatively characterising the corresponding transcriptome, and in
 CC detecting RNA transcripts and splice variants of human or animal
 CC transcriptomes. The libraries may also be used as specialised mini
 CC libraries to detect transcripts of a sub-transcriptome under a particular
 CC biological or pathological state, and so allowing the detection of tissue
 CC - and pathology-specific genes such as those genes only expressed in
 CC specific tissue under a specific pathological condition; to detect
 CC developmental specific genes; and to detect RNA transcripts and splice
 CC variants of a transcriptome of a patient suffering from a particular
 CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
 CC rats, humans and mice, which are used in the exemplification of the
 CC present invention. N.B. The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 60 BP; 11 A; 19 C; 16 G; 14 T; 0 U; 0 Other;

Query Match 55.7%; Score 15.6; DB 6; Length 60;

Best Local Similarity 81.8%; Pred. No. 3.1e+03;

Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CTGGAACCTGCATCCAAATTCAG 24

Db 46 CAGGAACTGACCCAGATTCCG 25

RESULT 15
ABN28998

ID ABN28998 standard; DNA; 65 BP.

XX AC ABN28998;

XX DT 15-JUL-2002 (first entry)

XX DE Rat spliced transcript detection oligonucleotide SEQ ID NO:1746.

XX KW Human; mouse; rat; splice transcript; detection; RNA transcript;

XX KW splice variant; transcriptome; oligonucleotide library; ss.

XX OS Rattus norvegicus.

XX PN WO200210449-A2.

XX PD 07-FEB-2002.

XX PF 20-JUL-2001; 2001WO-IB001903.

XX PR 28-JUL-2000; 2000US-0221607P.

XX PR 02-MAY-2001; 2001US-0287724P.

XX PA (COMP-) COMPUGEN INC.

XX PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;

XX DR WPI; 2002-257383/30.

XX PT New oligonucleotide libraries comprising oligonucleotides which
selectively hybridize to mRNAs transcribed from a transcription unit of a
genome, useful for detecting tissue-, pathology-, and developmental-
specific genes.

XX PS Example 1; SEQ ID NO 1746; 47pp; English.

XX CC The present invention describes oligonucleotide libraries for detecting
messenger RNAs that populate a (sub-)transcriptome, where the (sub-
)transcriptome comprises messenger RNAs transcribed from multiple
transcription units that populate a genome. The library comprises several
oligonucleotides, each capable of hybridizing selectively to a set of
messenger RNAs transcribed from a given transcription unit of the genome,
which encodes one or more messenger RNA splice variants. The
oligonucleotide libraries are useful for detecting mRNAs from a
biological sample, in expression profiling studies, in qualitatively or
quantitatively characterizing the corresponding transcriptome, and in
detecting RNA transcripts and splice variants of human or animal
transcriptomes. The libraries may also be used as specialised mini
libraries to detect transcripts of a sub-transcriptome under a particular
biological or pathological state, and so allowing the detection of tissue
- and pathology-specific genes such as those genes only expressed in
specific tissue under a specific pathological condition; to detect
developmental specific genes; and to detect RNA transcripts and splice
variants of a transcriptome of a patient suffering from a particular
disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
rats, humans and mice, which are used in the exemplification of the
present invention. N.B. The sequence data for this patent did not form
part of the printed specification, but was obtained in electronic format
directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 65 BP; 15 A; 17 C; 19 G; 14 T; 0 U; 0 Other;

Query Match 55.7%; Score 15.6; DB 6; Length 65;
Best Local Similarity 81.8%; Pred. No. 3 1e+03;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTCGAACTGACCCAGATTCCG 22

Db 23 GGCTGAACTGAAGCCACATTC 44

Search completed: September 1, 2004, 22:43:28
Job time : 234 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2004, 21:55:47 ; Search time 49 Seconds
(without alignments)
317.115 Million cell updates/sec

Title: US-10-028-415-2

Perfect score: 28
Sequence: 1 gctggaactgcattccaaattcagggttc 28

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 939290

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	100.0	28	2	US-08-713-557B-2
2	28	100.0	28	2	US-08-713-557B-14
3	28	100.0	50	2	US-08-713-557B-35
4	16	57.1	25	2	US-08-713-557B-29
5	16	57.1	36	2	US-08-860-174A-18
6	16	57.1	36	3	US-09-171-025-22
7	16	57.1	36	4	US-09-742-693-5
8	15	54.3	40	5	PCT-US95-11405-20
9	15	53.6	43	1	US-07-938-084-15
10	15	53.6	68	2	US-08-459-135A-3
11	15	53.6	68	2	US-08-459-135A-4
12	15	53.6	68	3	US-08-495-559-3
13	15	53.6	68	3	US-08-495-559-4
14	14.6	52.1	58	3	US-09-140-466-9
15	14.6	52.1	58	3	US-09-140-466-10
16	14.6	52.1	59	1	US-08-313-127A-2
17	14.6	52.1	66	1	US-08-313-127A-1
18	14.4	51.4	47	4	US-08-422-978-3861
19	14.4	51.4	59	4	US-08-956-171E-5133
20	14.4	51.4	80	3	US-09-077-690-3
21	14.2	50.7	31	1	US-08-347-826A-6
22	14.2	50.7	39	1	US-08-644-864B-35
23	14.2	50.7	39	2	US-08-761-277A-35
24	14	50.0	14	2	US-08-713-557B-12
25	14	50.0	14	2	US-08-713-557B-13
26	14	50.0	14	2	US-08-713-557B-15
27	14	50.0	14	2	US-08-713-557B-16

Sequence 17, Appl
Sequence 18, Appl
Sequence 4, Appl
Sequence 39, Appl
Sequence 39, Appl
Sequence 70, Appl
Sequence 132, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 17, Appl
Sequence 17, Appl
Sequence 257, Appl
Sequence 257, Appl
Sequence 64, Appl
Sequence 1620, Ap
Sequence 2877, Ap
Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-713-557B-2
; Sequence 2, Application US/08713557B
; Patent No. 5912189
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Rudert, Fritz
; TITLE OF INVENTION: CD95 REGULATORY GENE SEQUENCES
; TITLE OF INVENTION: AND TRANSCRIPTION FACTORS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/713,557B
; FILING DATE: 30-AUG-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Speckman, Ann W
; REGISTRATION NUMBER: 31,881
; REFERENCE/DOCKET NUMBER: 11000.1004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-713-557B-2

Query Match 100.0%; Score 28; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCTGGAAGTGCATCCAAATTCAGGTTC 28

Db 1 GTCTGGAAGTGCATCCAAATTCAGGTTTC 28
|||||
RESULT 2
US-08-713-557B-14/c
; Sequence 14, Application US/08713557B
; Patent No. 5912168
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Rudert, Fritz
; TITLE OF INVENTION: CD95 REGULATORY GENE SEQUENCES
; TITLE OF INVENTION: AND TRANSCRIPTION FACTORS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/713.557B
; FILING DATE: 30-AUG-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Speckman, Ann W
; REGISTRATION NUMBER: 31,881
; REFERENCE/DOCKET NUMBER: 11000.1004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-713-557B-14
Query Match 100.0%; Score 28; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTCTGGAAGTGCATCCAAATTCAGGTTTC 28
|||||
Db 28 GTCTGGAAGTGCATCCAAATTCAGGTTTC 1
|||||
RESULT 3
US-08-713-557B-35
; Sequence 35, Application US/08713557B
; Patent No. 5912168
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Rudert, Fritz
; TITLE OF INVENTION: CD95 REGULATORY GENE SEQUENCES
; TITLE OF INVENTION: AND TRANSCRIPTION FACTORS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA

COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/713.557B
FILING DATE: 30-AUG-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Speckman, Ann W
REGISTRATION NUMBER: 31,881
REFERENCE/DOCKET NUMBER: 11000.1004
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-713-557B-35
Query Match 100.0%; Score 28; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 0.0003;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTCTGGAAGTGCATCCAAATTCAGGTTTC 28
|||||
Db 1 GTCTGGAAGTGCATCCAAATTCAGGTTTC 28
|||||
RESULT 4
US-08-713-557B-29
; Sequence 29, Application US/08713557B
; Patent No. 5912168
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Rudert, Fritz
; TITLE OF INVENTION: CD95 REGULATORY GENE SEQUENCES
; TITLE OF INVENTION: AND TRANSCRIPTION FACTORS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/713.557B
; FILING DATE: 30-AUG-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Speckman, Ann W
; REGISTRATION NUMBER: 31,881
; REFERENCE/DOCKET NUMBER: 11000.1004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-713-557B-29

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:

INFORMATION FOR SEQ ID NO: 29:

SEQUENCE CHARACTERISTICS:
LENGTH: 25 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-713-557B-29

Query Match 57.1%; Score 16; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCTGGAAGTCATCC 16
|||||
DB 10 GTCTGGAAGTCATCC 25

RESULT 5

US-08-860-174A-18
Sequence 18, Application US/08860174A
Patent No. 5989830

GENERAL INFORMATION:

APPLICANT: DAVIS, Paul James
APPLICANT: VAN DER LOGT, Cornelis Paul Erik
APPLICANT: VERHOEIJEN, Martine Elisa
APPLICANT: WILSON, Steve
TITLE OF INVENTION: A BIFUNCTIONAL OR BIVALENT
TITLE OF INVENTION: ANTIBODY FRAGMENT ANALOGUE
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY, MADISON & SUTRO LLP
STREET: 1100 New York Avenue, N.W.
STREET: 9th Floor, East Tower
CITY: WASHINGTON, D.C.
STATE:

COUNTRY: UNITED STATES

ZIP: 20005-3918

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS DOS Text
SOFTWARE: MS WORD, WINDOWS 95, Ver. 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/860,174A
FILING DATE: June 16, 1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: EP 95307332.7
FILING DATE: October 16, 1995
APPLICATION NUMBER: PCT/EP/96/03605
FILING DATE: August 14, 1996

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:
LENGTH: 36 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid

DESCRIPTION: /desc = "synthetic DNA"

IMMEDIATE SOURCES:

CLONE: primer DBL.5

US-08-860-174A-18

Query Match 57.1%; Score 16; DB 2; Length 36;
Best Local Similarity 79.2%; Pred. No. 1.3e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 GGAACTGCATCCAAATTCAGGTTTC 28
|||||
DB 13 GGAGCTGCATGCAAAATTCATTTC 36

RESULT 6

US-09-171-025-22

Sequence 22, Application US/09171025
Patent No. 6239259

GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: Multivalent and multispecific antigen-binding
TITLE OF INVENTION: protein
NUMBER OF SEQUENCES: 27
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,025
FILING DATE:

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 36 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "synthetic DNA"
IMMEDIATE SOURCE:
CLONE: primer DBL.10

US-09-171-025-22

Query Match 57.1%; Score 16; DB 3; Length 36;
Best Local Similarity 79.2%; Pred. No. 1.3e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 GGAACTGCATCCAAATTCAGGTTTC 28
|||||
DB 13 GGAGCTGCATGCAAAATTCATTTC 36

RESULT 7

US-09-742-693-5

Sequence 5, Application US/09742693
Patent No. 6579842

GENERAL INFORMATION:

APPLICANT: HOWELL, Steven
APPLICANT: LITTLE, Julie C.
APPLICANT: VAN DER LOGT, Cornelis P.
APPLICANT: PAREY, Neil J.
TITLE OF INVENTION: METHOD OF TREATING FABRICS
FILE REFERENCE: C7536 (V)
CURRENT APPLICATION NUMBER: US/09/742,693
CURRENT FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: EP99310431.4
PRIOR FILING DATE: 1999-12-22
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5

LENGTH: 36

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence:Primer

US-09-742-693-5

Query Match 57.1%; Score 16; DB 4; Length 36;
Best Local Similarity 79.2%; Pred. No. 1.3e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 GGAACTGCATCCAAATTCAGGTTTC 28
|||||
DB 13 GGAGCTGCATGCAAAATTCATTTC 36

RESULT 8
PCT-US95-11405-20
; Sequence 20, Application PC/TUS95111405
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PHOSPHOROUS-32 LABELING OF ANTIBODIES
; TITLE OF INVENTION: FOR CANCER THERAPY
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/11405
; FILING DATE: 18-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/308,103
; FILING DATE: 16-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/599/IWIN
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
PCT-US95-11405-20

Query Match 54.3%; Score 15.2; DB 5; Length 40;
Best Local Similarity 71.4%; Pred. No. 3.2e+02;
Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GTCTGGAAGTGCATCCAAATTCAGGTTTC 28
Db 11 GCCGAGATCTGCATCCAAATTCATTTTC 38

RESULT 9
US-07-938-084-15
; Sequence 15, Application US/07938084
; Patent No. 5464945
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Rebecca L.
; APPLICANT: Walsh, P. Sean
; TITLE OF INVENTION: A Chemiluminescent Method for the
; TITLE OF INVENTION: Quantitation of Human DNA
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: NJ
; COUNTRY: U.S.A.
; ZIP: 07110-1199
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/938,084
; FILING DATE: 19920828
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sias Ph.D., Stacey R.
; REGISTRATION NUMBER: 32,630
; REFERENCE/DOCKET NUMBER: 8669
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 814-2863
; TELEFAX: (510) 522-1285
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-938-084-15

Query Match 53.6%; Score 15; DB 1; Length 43;
Best Local Similarity 78.3%; Pred. No. 4e+02;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 6 GAATGTCATCCAAATTCAGGTTTC 28
Db 1 GAATGTCATCCGACATCTGGTTC 23

RESULT 10
US-08-459-135A-3
; Sequence 3, Application US/08459135A
; Patent No. 5955078
; GENERAL INFORMATION:
; APPLICANT: BURNHAM, Martin Karl Russel
; APPLICANT: CHOPEA, Ian
; APPLICANT: CRITCHLEY, Ian Alfred
; APPLICANT: KNOWLES, David Justin Charles
; TITLE OF INVENTION: FIBRONECTIN BINDING PROTEIN: MONOCLONAL ANTIBODY AND THEIR US
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fastseq, Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,135A
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB/94/00215
; FILING DATE: 04-Feb-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R.
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P30591C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 base pairs
; TYPE: nucleic acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: GENOMIC DNA
US-08-459-135A-3

Query Match      53.6%; Score 15; DB 2; Length 68;
Best Local Similarity 78.3%; Pred. No. 4.5e+02;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 TCTGGAAGTGCATCCCAATTTCAG 24
    |||||
Db 46 TCTGGATCCGCATCGCAATTCGG 68

RESULT 11
US-08-459-135A-4
; Sequence 4, Application US/08459135A
; Patent No. 5955078
; GENERAL INFORMATION:
; APPLICANT: BURNHAM, Martin Karl Russel
; APPLICANT: CHOPRA, Ian
; APPLICANT: CRITCHLEY, Ian Alfred
; APPLICANT: KNOWLES, David Justin Charles
; TITLE OF INVENTION: FIBRONECTIN BINDING PROTEIN: MONOCLONAL ANTIBODY AND THEIR USE IN PREVENTING BACTERIAL ADHESION
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fastseq, Version 2.0
; CURRENT APPLICATION DATA: US/08/459,135A
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: PCT/GB/94/00215
; FILING DATE: 04-Feb-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 36,891
; REFERENCE/DOCKET NUMBER: P30591C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: GENOMIC DNA
US-08-459-135A-4

Query Match      53.6%; Score 15; DB 2; Length 68;
Best Local Similarity 78.3%; Pred. No. 4.5e+02;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 TCTGGAAGTGCATCCCAATTTCAG 24
    |||||
Db 46 TCTGGATCCGCATCGCAATTCGG 68

RESULT 12
US-08-495-559-3
; Sequence 3, Application US/08495559A
; Patent No. 6054572
; GENERAL INFORMATION:
; APPLICANT: BURNAM, Martin Karl Russel
; APPLICANT: CHOPRA, Ian
; APPLICANT: CRITCHLEY, Ian Alfred
; APPLICANT: KNOWLES, David Justin Charles
; TITLE OF INVENTION: FIBRONECTIN BINDING PROTEIN: MONOCLONAL ANTIBODY AND THEIR USE IN PREVENTING BACTERIAL ADHESION
; FILE REFERENCE: P30591
; CURRENT APPLICATION NUMBER: US/08/495,559A
; CURRENT FILING DATE: 1995-08-03
; EARLIER APPLICATION NUMBER: GB 9302289.5
; EARLIER FILING DATE: 1993-02-05
; EARLIER APPLICATION NUMBER: GB 9321592.9
; EARLIER FILING DATE: 1993-10-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 68
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-08-495-559-3

Query Match      53.6%; Score 15; DB 3; Length 68;
Best Local Similarity 78.3%; Pred. No. 4.5e+02;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 TCTGGAAGTGCATCCCAATTTCAG 24
    |||||
Db 46 TCTGGATCCGCATCGCAATTCGG 68

RESULT 13
US-08-495-559-4
; Sequence 4, Application US/08495559A
; Patent No. 6054572
; GENERAL INFORMATION:
; APPLICANT: BURNAM, Martin Karl Russel
; APPLICANT: CHOPRA, Ian
; APPLICANT: CRITCHLEY, Ian Alfred
; APPLICANT: KNOWLES, David Justin Charles
; TITLE OF INVENTION: FIBRONECTIN BINDING PROTEIN: MONOCLONAL ANTIBODY AND THEIR USE IN PREVENTING BACTERIAL ADHESION
; FILE REFERENCE: P30591
; CURRENT APPLICATION NUMBER: US/08/495,559A
; CURRENT FILING DATE: 1995-08-03
; EARLIER APPLICATION NUMBER: GB 9302289.5
; EARLIER FILING DATE: 1993-02-05
; EARLIER APPLICATION NUMBER: GB 9321592.9
; EARLIER FILING DATE: 1993-10-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 68
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-08-495-559-4

Query Match      53.6%; Score 15; DB 3; Length 68;
Best Local Similarity 78.3%; Pred. No. 4.5e+02;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 TCTGGAAGTGCATCCCAATTTCAG 24
    |||||
Db 46 TCTGGATCCGCATCGCAATTCGG 68

RESULT 14
US-09-140-466-9/c
; Sequence 9, Application US/09140466
; Patent No. 6268160
; GENERAL INFORMATION:

```

; APPLICANT: CLOUGH, BARBARA
; APPLICANT: PREISER, PETER
; APPLICANT: WILSON, ROBERT
; TITLE OF INVENTION: AN EF-TU PROTEIN ENCODED ON THE PLASTID DNA OF THE
; TITLE OF INVENTION: MALARIA PARASITE AND PROTEIN SYNTHESIS INHIBITORS
; TITLE OF INVENTION: EFFECTIVE AS ANTI-MALARIAL COMPOUNDS
; FILE REFERENCE: N68837B GCW PJC DP
; CURRENT APPLICATION NUMBER: US/09/140,466
; CURRENT FILING DATE: 1998-08-26
; EARLIER APPLICATION NUMBER: US 60/056,246
; EARLIER FILING DATE: 1997-08-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 58
; TYPE: RNA
; ORGANISM: Plasmodium falciparum
US-09-140-466-9

Query Match 52.1%; Score 14.6; DB 3; Length 58;
Best Local Similarity 81.0%; Pred. No. 6.7e+02;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 7 AACTGCATCCAAATTCAGTT 27
|||||
DB 26 AACTGCTTCCAAATTACTTT 6

RESULT 15
US-09-140-466-10/c
; Sequence 10, Application US/09140466
; Patent No. 6268160
; GENERAL INFORMATION:
; APPLICANT: CLOUGH, BARBARA
; APPLICANT: PREISER, PETER
; APPLICANT: WILSON, ROBERT
; TITLE OF INVENTION: AN EF-TU PROTEIN ENCODED ON THE PLASTID DNA OF THE
; TITLE OF INVENTION: MALARIA PARASITE AND PROTEIN SYNTHESIS INHIBITORS
; TITLE OF INVENTION: EFFECTIVE AS ANTI-MALARIAL COMPOUNDS
; FILE REFERENCE: N68837B GCW PJC DP
; CURRENT APPLICATION NUMBER: US/09/140,466
; CURRENT FILING DATE: 1998-08-26
; EARLIER APPLICATION NUMBER: US 60/056,246
; EARLIER FILING DATE: 1997-08-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 58
; TYPE: RNA
; ORGANISM: Plasmodium falciparum
US-09-140-466-10

Query Match 52.1%; Score 14.6; DB 3; Length 58;
Best Local Similarity 81.0%; Pred. No. 6.7e+02;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 7 AACTGCATCCAAATTCAGTT 27
|||||
DB 26 AACTGCTTCCAAATTACTTT 6

Search completed: September 1, 2004, 23:07:49
Job time : 50 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2004, 22:39:33 ; Search time 242 Seconds
(without alignments)
569.421 Million cell updates/sec

Title: US-10-028-415-2
Perfect score: 28
Sequence: 1 gctgggaactgcataaattcagggttc 28

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 323270 seqs, 2460713050 residues
Total number of hits satisfying chosen parameters: 2018620

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications NA:*
- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
 - 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
 - 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
 - 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
 - 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
 - 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
 - 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
 - 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
 - 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
 - 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
 - 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
 - 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
 - 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
 - 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
 - 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
 - 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
 - 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
 - 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
 - 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	28	100.0	28	10	US-09-997-905A-29
C 2	28	100.0	28	14	US-10-028-415-2
C 3	28	100.0	28	14	US-10-028-415-11
C 4	17.2	61.4	65	10	US-09-908-975-28733
C 5	16.6	59.3	60	10	US-09-908-975-9817
C 6	16	57.1	36	9	US-09-742-693-5
C 7	16	57.1	36	15	US-10-112-261B-10
C 8	15.8	56.4	35	15	US-10-131-827-4796
C 9	15.6	55.7	60	10	US-09-908-975-12681
C 10	15.6	55.7	60	10	US-09-908-975-1746
C 11	15.4	55.0	65	10	US-09-908-975-7331
C 12	15.2	54.3	65	15	US-10-032-585-3449
C 13	15.2	54.3	84	15	US-10-032-750-212
C 14	15	53.6	24	13	US-10-308-264-697

Sequence 12, Appl
Sequence 1855, A
Sequence 19479, A
Sequence 20568, A
Sequence 3049, Ap
Sequence 5559, Ap
Sequence 9, Appl
Sequence 10, Appl
Sequence 2841, Ap
Sequence 5, Appl
Sequence 3861, Ap
Sequence 5133, Ap
Sequence 5133, Ap
Sequence 6927, Ap
Sequence 18337, A
Sequence 32090, A
Sequence 32302, A
Sequence 187, App
Sequence 187, App
Sequence 184, App
Sequence 35, Appl
Sequence 35, Appl
Sequence 187, App
Sequence 490, App
Sequence 18477, A
Sequence 23008, A
Sequence 29773, A
Sequence 23913, A
Sequence 9, Appl
Sequence 10, Appl
Sequence 12, Appl

27 10 US-09-997-905A-12
60 10 US-09-908-975-18855
60 10 US-09-908-975-19479
60 10 US-09-908-975-20568
41 12 US-10-035-833A-3049
41 12 US-10-035-833A-5559
58 9 US-09-845-335-9
58 9 US-09-845-335-10
65 15 US-10-032-585-2841
30 15 US-10-092-263-5
47 16 US-10-349-143-3861
59 8 US-08-781-986A-5133
59 13 US-10-329-624-5133
60 10 US-09-908-975-6927
60 10 US-09-908-975-18337
60 10 US-09-908-975-32090
60 10 US-09-908-975-32302
60 16 US-10-289-695-187
22 15 US-10-410-998-187
22 15 US-10-005-338B-184
39 9 US-09-925-664-35
39 12 US-09-925-182-35
54 13 US-09-837-306-187
54 16 US-10-045-674-490
60 10 US-09-908-975-18477
60 10 US-09-908-975-23008
65 10 US-09-908-975-29773
88 9 US-09-864-761-23913
14 14 US-10-028-415-9
14 14 US-10-028-415-10
14 14 US-10-028-415-12

ALIGNMENTS

RESULT 1
US-09-997-905A-29/c
; Sequence 29, Application US/09997905A
; Publication No. US20030074684A1
; GENERAL INFORMATION:
; APPLICANT: Benitec Australia Ltd
; APPLICANT: State of Queensland through its Department of Primary Industries
; TITLE OF INVENTION: Control of Gene Expression
; FILE REFERENCE: M80219470
; CURRENT APPLICATION NUMBER: US/09/997,905A
; CURRENT FILING DATE: 2002-11-30
; PRIOR APPLICATION NUMBER: US 09/100,812
; PRIOR FILING DATE: 1998-06-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 28
; TYPE: DNA
; ORGANISM: double-stranded
US-09-997-905A-29

Query Match 100.0%; Score 28; DB 10; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTGGAACTGCATCCAAATTCAGGTTC 28
Db 28 GCTGGAACTGCATCCAAATTCAGGTTC 1

RESULT 2
US-10-028-415-2
; Sequence 2, Application US/10028415
; Publication No. US20020151063A1
; GENERAL INFORMATION:
; APPLICANT: Lasham, Annette
; APPLICANT: Watson, James D.

```

; TITLE OF INVENTION: Methods for Modulating Apoptotic Cell
; TITLE OF INVENTION: Death
; FILE REFERENCE: 11000.1004c3
; CURRENT APPLICATION NUMBER: US/10/028,415
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: PCT/NZ01/00286
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: US 09/724,809
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/036,004
; PRIOR FILING DATE: 1998-03-04
; PRIOR APPLICATION NUMBER: US 08/713,557
; PRIOR FILING DATE: 1996-08-30
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 28
; TYPE: DNA
; ORGANISM: Human
US-10-028-415-2

```

```

Query Match 100.0%; Score 28; DB 14; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GTCGGAAGTGCATCCAAATTCAGGTTTC 28
DB 1 GTCGGAAGTGCATCCAAATTCAGGTTTC 28

```

RESULT 3

```

US-10-028-415-11/c
; Sequence 11, Application US/10028415
; Publication No. US20020151063A1
; GENERAL INFORMATION:
; APPLICANT: Lasham, James D.
; APPLICANT: Watson, Annette
; TITLE OF INVENTION: Methods for Modulating Apoptotic Cell
; TITLE OF INVENTION: Death
; FILE REFERENCE: 11000.1004c3
; CURRENT APPLICATION NUMBER: US/10/028,415
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: PCT/NZ01/00286
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: US 09/724,809
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/036,004
; PRIOR FILING DATE: 1998-03-04
; PRIOR APPLICATION NUMBER: US 08/713,557
; PRIOR FILING DATE: 1996-08-30
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 28
; TYPE: DNA
; ORGANISM: Human
US-10-028-415-11

```

```

Query Match 100.0%; Score 28; DB 14; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GTCGGAAGTGCATCCAAATTCAGGTTTC 28
DB 28 GTCGGAAGTGCATCCAAATTCAGGTTTC 1

```

RESULT 4

```

US-09-908-975-28733/c
; Sequence 28733, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi

```

```

; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 28733
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-28733

```

```

Query Match 61.4%; Score 17.2; DB 10; Length 65;
Best Local Similarity 86.4%; Pred. No. 4.5e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 3 CTGGAAGTGCATCCAAATTCAG 24
DB 61 CTGCACTGCGAGCCAAATTCAG 40

```

RESULT 5

```

US-09-908-975-9817/c
; Sequence 9817, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9817
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-9817

```

```

Query Match 59.3%; Score 16.6; DB 10; Length 60;
Best Local Similarity 82.6%; Pred. No. 8.5e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 3 CTGGAAGTGCATCCAAATTCAGG 25
DB 35 CTGTCCTGCGTCCAAAGTCAGG 13

```

RESULT 6

```

US-09-742-693-5
; Sequence 5, Application US/09742693
; Patent No. US20020019324A1
; GENERAL INFORMATION:
; APPLICANT: HOWELL, Steven
; APPLICANT: LITTLE, Julie C.

```

```
; APPLICANT: VAN DER LOGT, Cornelis P.
; APPLICANT: PARRY, Neil J.
; TITLE OF INVENTION: METHOD OF TREATING FABRICS
; FILE REFERENCE: C7536 (V)
; CURRENT APPLICATION NUMBER: US/09/742,693
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: EP99310431.4
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 36
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Primer
US-09-742-693-5

Query Match      57.1%; Score 16; DB 9; Length 36;
Best Local Similarity 79.2%; Pred. No. 1.5e+03;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 GGAAGTGCATCCAAATTCAGGTTTC 28
   ||||| ||||| ||||| ||||| |||||
Db 13 GGAGCTGCATGCAAAATTCATTTC 36

RESULT 7
US-10-112-261B-10
; Sequence 10, Application US/10112261B
; Publication No. US20030100570A1
; GENERAL INFORMATION:
; APPLICANT: Jameson, Bradford A.
; APPLICANT: Tretiakova, Anna A.
; APPLICANT: Albert, Ross
; APPLICANT: Davidson, Harold Carter
; TITLE OF INVENTION: IMMUNOMODULATION AND EFFECT ON CELL PROCESSES RELATING TO SEROTONIN
; FILE REFERENCE: 054480-5001-00
; CURRENT APPLICATION NUMBER: US/10/112,261B
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: US 60/353,883
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/345,295
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 60/280,296
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 36
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-112-261B-10

Query Match      57.1%; Score 16; DB 15; Length 36;
Best Local Similarity 79.2%; Pred. No. 1.5e+03;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 TCTGGAAGTGCATCCAAATTCAGG 25
   ||||| ||||| ||||| ||||| |||||
Db 10 TCAGCACTACTCTCCAAATTCAGG 33

RESULT 8
US-10-131-827-4796
; Sequence 4796, Application US/10131827
; Publication No. US20040009479A1
; GENERAL INFORMATION:
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
```

```
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUN
; FILE REFERENCE: 506612000120
; CURRENT APPLICATION NUMBER: US/10/131,827
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: US 10/006,290
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/296,764
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 9090
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4796
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-827-4796

Query Match      56.4%; Score 15.8; DB 16; Length 50;
Best Local Similarity 74.1%; Pred. No. 1.9e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 TCTGGAAGTGCATCCAAATTCAGGTTTC 28
   ||||| ||||| ||||| ||||| |||||
Db 24 TCTGTTACTAGCTACAAATTCGGTTTC 50

RESULT 9
US-09-908-975-12681/c
; Sequence 12681, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12681
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-12681

Query Match      55.7%; Score 15.6; DB 10; Length 60;
Best Local Similarity 81.8%; Pred. No. 2.5e+03;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CTGGAAGTGCATCCAAATTCAG 24
   ||||| ||||| ||||| ||||| |||||
Db 46 CAGGAAGTGCACCCAGATTCGG 25

RESULT 10
US-09-908-975-1746
; Sequence 1746, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
```

; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1746
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-908-975-1746

Query Match 55.7%; Score 15.6; DB 10; Length 65;
Best Local Similarity 81.8%; Pred. No. 2.5e+03;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCTGGAAGTGCATCCAAATTC 22
| | | | | | | | | | | | | | | | | | | | | |
DB 23 GGCTGGAAGTGAAGCCAAATTC 44

RESULT 11
US-09-908-975-7331/c
; Sequence 7331, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICING
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7331
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-7331

Query Match 55.0%; Score 15.4; DB 10; Length 60;
Best Local Similarity 76.0%; Pred. No. 3.1e+03;
Matches 19; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 CTGGAAGTGCATCCAAATTCAGGTT 27
| | | | | | | | | | | | | | | | | | | | | |
DB 47 CTGGAGTGCACCAACTCAGGTT 23

RESULT 12
US-10-032-585-3449
; Sequence 3449, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585

; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3449
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Cardida albicans
US-10-032-585-3449

Query Match 54.3%; Score 15.2; DB 15; Length 65;
Best Local Similarity 85.0%; Pred. No. 3.9e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 TGGAACTGCATCCAAATTC 23
| | | | | | | | | | | | | | | | | | | | | |
DB 18 TGGACCAGCATCCAAATCA 37

RESULT 13
US-10-092-750-212
; Sequence 212, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 212
; LENGTH: 84
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 147
; OTHER INFORMATION: n = A,T,C or G
US-10-092-750-212

Query Match 54.3%; Score 15.2; DB 15; Length 84;
Best Local Similarity 85.0%; Pred. No. 4.1e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 TGGAACTGCATCCAAATTC 23
| | | | | | | | | | | | | | | | | | | | | |
DB 34 TGGAACTGCATCCAAACTAA 53

RESULT 14
US-10-308-264-697/c
; Sequence 697, Application US/10308264
; Publication No. US20040029133A1
; GENERAL INFORMATION:
; APPLICANT: Herrnstadt, Corinna
; TITLE OF INVENTION: MITOCHONDRIAL DNA POLYMORPHISM
; FILE REFERENCE: 660088.461
; CURRENT APPLICATION NUMBER: US/10/308,264
; CURRENT FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 697
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 697
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer
US-10-308-264-697

Query Match 53.6%; Score 15; DB 13; Length 24;
Best Local Similarity 78.3%; Pred. No. 4e+03;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 6 GAATGCTCCAAATTCAGGTTT 28
Db 24 GAATGCTATCCGACATCTGTTT 2

RESULT 15
US-09-997-905A-12/c
; Sequence 12, Application US/09997905A
; Publication No. US20030074684A1
; GENERAL INFORMATION:
; APPLICANT: Benitec Australia Ltd
; TITLE OF INVENTION: State of Queensland through its Department of Primary Industries
; FILE REFERENCE: M80219470
; CURRENT APPLICATION NUMBER: US/09/997,905A
; CURRENT FILING DATE: 2002-11-30
; PRIOR APPLICATION NUMBER: US 09/100,812
; PRIOR FILING DATE: 1998-06-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 27
; TYPE: DNA
; ORGANISM: virus
US-09-997-905A-12

Query Match 52.9%; Score 14.8; DB 10; Length 27;
Best Local Similarity 73.1%; Pred. No. 5.1e+03;
Matches 19; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 TCTGGAACTGCTCCAAATTCAGGTT 27
Db 26 TCTGGAACTGGGTAGAGTTACGTT 1

Search completed: September 1, 2004, 23:42:42
Job time : 244 secs

This Page Blank (uspt.)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2004, 22:35:18 ; Search time 1765 Seconds
(without alignments)
473.734 Million cell updates/sec

Title: US-10-028-415-2

Perfect score: 28

Sequence: 1 gctggaactgcatccaaattcaggttc 28

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 569298

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

EST.*
1: em_estba:*
2: em_esthum:*
3: em_estnu:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_estc:*
10: gb_estc2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_lam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	18.4	65.7	67	14	CD881233
2	16.6	59.3	85	9	AU265805
3	16	57.1	52	10	AW396361
4	16	57.1	64	10	BE638254

C 5	15.8	56.4	73	13	BQ086192
C 6	15.8	56.4	73	29	CG532825
C 7	15.8	56.4	91	9	AA814152
C 8	15.8	56.4	100	28	AZ761676
C 9	15.6	55.7	66	14	CD946266
C 10	15.6	55.7	67	14	CD946207
C 11	15.2	54.3	50	9	AU105882
C 12	15.2	54.3	56	10	AW268320
C 13	15.2	54.3	59	12	BJ041364
C 14	15.2	54.3	61	9	AI174615
C 15	15.2	54.3	76	9	AA671040
C 16	15.2	54.3	88	9	AI000664
C 17	15.2	54.3	89	12	BI490973
C 18	15.2	54.3	95	9	AI855768
C 19	15.2	54.3	98	10	BE138707
C 20	15	53.6	50	9	AU102350
C 21	15	53.6	50	9	AU102351
C 22	15	53.6	50	9	AU105884
C 23	15	53.6	84	29	CG564129
C 24	15	53.6	92	14	CA803412
C 25	15	53.6	94	29	CG646872
C 26	14.8	52.9	45	28	BH850369
C 27	14.8	52.9	58	9	AL805765
C 28	14.8	52.9	60	14	D45451
C 29	14.8	52.9	61	28	BH863711
C 30	14.8	52.9	77	9	AA980728
C 31	14.8	52.9	82	29	CNS041CQ
C 32	14.8	52.9	85	12	BI407547
C 33	14.8	52.9	94	14	N40508
C 34	14.8	52.9	98	28	CC019463
C 35	14.6	52.1	59	10	BG041906
C 36	14.6	52.1	95	14	CB187735
C 37	14.4	51.4	67	9	AA654686
C 38	14.4	51.4	79	13	BU637417
C 39	14.4	51.4	87	14	Z20306
C 40	14.4	51.4	100	14	T97877
C 41	14.2	50.7	43	28	AZ803705
C 42	14.2	50.7	50	9	AU104942
C 43	14.2	50.7	54	13	BX551631
C 44	14.2	50.7	58	9	AL659070
C 45	14.2	50.7	61	29	CNS02YTI

ALIGNMENTS

RESULT 1
CD881233
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CD881233 67 bp mRNA linear EST 14-JUL-2003
F1.102G08F010328 F1 Triticum aestivum cDNA clone F1102G08, mRNA
sequence.

CD881233.1 GI:32640120

EST.
Triticum aestivum (bread wheat)
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 67)

Genoplante.
Genoplante, a major partnership french program in plant genomics
Unpublished (2003)
Contact: Genoplante
Genoplante

93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10

This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (<http://www.genoplante.com>)
and <http://genoplante-info.infobiogen.fr>.

Location/Qualifiers
1..67

FEATURES
source

ORIGIN

```

/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="F1102G08"
/tissue_type="leaf one"
/clone_lib="F1"

```

Query Match 65.7%; Score 18.4; DB 14; Length 67;
 Best Local Similarity 78.6%; Pred. No. 2e+03;
 Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GTCTGGAACTGCATCCAAATTCAGGTTTC 28
 |||||
 Db 32 GTCTGGAACTGAAACTAAATTAAGGTGC 59

RESULT 2
 AU265805 85 bp mRNA linear EST 10-MAY-2002
 LOCUS AU265805 VS Dictyostelium discoideum cDNA clone VSF794 3', mRNA
 DEFINITION

ACCESSION AU265805
 VERSION AU265805
 KEYWORDS EST.
 SOURCE Dictyostelium discoideum

ORGANISM Dictyostelium discoideum

REFERENCE 1 (bases 1 to 85)
 AUTHORS Urushihara, H., Morio, T., Saito, T., Kori, E., Ochiai, H., Maeda, M.,
 Takeuchi, I., Kohara, Y. and Tanaka, Y.

TITLE Population analysis of cDNAs from unicellular and multicellular
 stages of Dictyostelium discoideum

JOURNAL Unpublished (2002)
 COMMENT Contact: Hideko Urushihara
 Institute of Biological Sciences
 University of Tsukuba
 1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
 Tel: 81-298-53-4664
 Fax: 81-298-53-6614
 Email: hideko@biol.tsukuba.ac.jp.

FEATURES
 source
 1..85
 /organism="Dictyostelium discoideum"
 /mol_type="mRNA"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="VSF794"
 /sex="mat A"
 /dev_stage="vegetative"
 /clone_lib="VS"

ORIGIN

```

Query Match 59.3%; Score 16.6; DB 9; Length 85;  

Best Local Similarity 79.2%; Pred. No. 1.2e+04;  

Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY 5 GGAAGTGCATCCAAATTCAGGTTTC 28
 |||||
 Db 39 GGAGTTCCTTACAAATTCAGGTTTC 62

RESULT 3
 AW396361/c 52 bp mRNA linear EST 03-DEC-2001
 LOCUS sh27b06.y1 Gm-cl016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 DEFINITION Gm-cl016-5964 5', mRNA sequence.

ACCESSION AW396361
 VERSION AW396361
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max

REFERENCE

AUTHORS

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.

1 (bases 1 to 52)

Shoemaker, R., Keim, P., Vodka, L., Erpelting, J., Corryell, V.,
 Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
 Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
 Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
 Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
 McCann, R., Waterston, R. and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information

call: (800)-533-4363 or contact via email: ccu@resgen.com

Insert Length: 252 Std Error: 0.00

Seq primer: -40RP from Gibco.

Location/Qualifiers

1..52

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl016-5964"

/tissue_type="immature flowers of field grown plants"

/lab_host="XL10-Gold"

/clone_lib="Gm-cl016"

/note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2:

XhoI; This cDNA library was constructed from mRNA isolated

from immature flowers of field grown plants. The cDNA

library was prepared using the Stratagene pBluescript II

XR library construction kit. Complementary DNA was

synthesized from mRNA using a primer consisting of a poly

(dT) sequence with a XhoI restriction site. EcoRI adapters

were ligated to the blunt-ended cDNA fragments followed by

XhoI digestion. The cDNA fragments were directionally

cloned into the EcoRI-XhoI restriction site of the

pBluescript vector. The ligated cDNA fragments were

transformed into XL10-Gold host cells. This library was

constructed by Dr. Randy Shoemaker and Dr. John

Erpelting."

ORIGIN

Query Match 57.1%; Score 16; DB 10; Length 52;
 Best Local Similarity 79.2%; Pred. No. 1.8e+04;
 Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CTGGAACTGCATCCAAATTCAGGT 26

|||||
 Db 50 CTCTAACTGAATCCCAATTCAGTT 27

RESULT 4

BE638254

LOCUS BE638254

DEFINITION

BE638254.1 GI:9931997

KEYWORDS EST.

SOURCE Onchocerca volvulus

ORGANISM Onchocerca volvulus

Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;

Onchocercidae; Onchocerca.

REFERENCE 1 (bases 1 to 64)

BE638254 64 bp mRNA linear EST 28-AUG-2000

SWOVMFCAR17A08SK Onchocerca volvulus microfilaria cDNA

(SAW98MLW-Ovml) Onchocerca volvulus cDNA clone SOWMFCAR17A08 5',

mRNA sequence.

AUTHORS
TITLE Genes expressed in microfilaria of Onchocerca volvulus
JOURNAL Unpublished (1999)
COMMENT Contact: Steven A. Williams
 Molecular Parasitology
 Smith College Department of Biological Sciences
 Department of Biological Sciences, Clark Science Center, Smith
 College, Northampton, MA, 01063, USA
 Tel: 4135853826
 Fax: 4135853786
 Email: genome@smith.edu
 Seq primer: pBluescript SK.
 Location/Qualifiers

FEATURES
 source
 1..64
 /organism="Onchocerca volvulus"
 /mol_type="mRNA"
 /db_xref="taxon:6282"
 /clone="SVOVMFCAR17A08"
 /dev_stage="microfilaria"
 /lab_host="X11-Blue MRP"
 /clone_lib="Onchocerca volvulus microfilaria cDNA
 (SAW98VLW-OvMF)"
 /note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
 Xho I; Filarial nematode parasite of humans. mRNA was
 prepared from approximately 200,000 microfilariae isolated
 from the skin of infected individuals from Kumba,
 Cameroon and converted to double-stranded cDNA using
 reverse transcriptase and oligo(dT) followed by RNase H
 and DNA pol I. The library has 7.8 x 10E4 independent
 recombinants and the average insert size is approximately
 1kb. The library was constructed by Michelle
 Lizotte-Waniewski. The library is available from
 Dr S.A.Williams, email: genome@smith.edu."

ORIGIN
 Query Match 57.1%; Score 16; DB 10; Length 64;
 Best Local Similarity 79.2%; Pred. No. 2e+04;
 Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 GGAAGTGCATCCAAATTCAGGTTTC 28
 ||||| ||||| ||||| ||||| |||||
 Db 8 GGAAGTGCATCCAAATTCAGGTTTC 31

RESULT 5
 BQ086192/c
 LOCUS BQ086192 73 bp mRNA linear EST 29-APR-2002
 DEFINITION i19f02.y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
 cDNA clone IMAGE:6135291 5', mRNA sequence.
 ACCESSION BQ086192
 VERSION BQ086192.1 GI:20045396
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 73)
 AUTHORS Melton,D., Brown,J., Keny,G., Permutt,A., Lee,C., Kaestner,K.,
 Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
 Hallier,I., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
 Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
 Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarishvili,R.,
 Williams,T., Jackson,Y. and Bowers,Y.
 Endocrine Pancreas Consortium
 Unpublished (2000)
 CONTACT: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Douglas Melton DNA sequencing by:
 Washington University Genome Sequencing Center This clone is
 available royalty-free through LBNL; please contact the IMAGE
 consortium (info@image.llnl.gov) for further information
 Putative full length read
 vector to vector length is 74
 Seq primer: -40RP from Gibco.
 Location/Qualifiers

FEATURES
 source
 1..73
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6135291"
 /sex="Both"
 /tissue_type="Islets of Langerhans"
 /dev_stage="Adult"
 /lab_host="DH10B"
 /clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
 /note="Organ: Pancreas; Vector: pSPOR1; Site_1: Not 1;
 Site_2: Sal 1; Starting library constructed using
 SuperScript Plasmid Library kit (Life Technologies). cDNA
 made by oligo-dT priming. Size-selected by column
 fractionation; average insert size 1.08 kb. Library was
 amplified once on solid support and plasmid DNA from
 library was prepared. The library DNA was normalized by
 method #4 from Bonaldo, Lennon, and Soares 1996 Genome
 Research 6:791-806; 0.5 microgram single-stranded library
 plasmid DNA was mixed with 5 micrograms PCR product
 representing library inserts and hybridized to an Ecot of
 20. Single-stranded (unhybridized) plasmids were isolated
 by hydroxyapatite chromatography and used to make this
 library."

ORIGIN
 Query Match 56.4%; Score 15.8; DB 13; Length 73;
 Best Local Similarity 74.1%; Pred. No. 2.5e+04;
 Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 TGTGAACTGCATCCAAATTCAGGTTTC 28
 ||||| ||||| ||||| ||||| |||||
 Db 30 TGGGGAAGTAAACCCCAAGGCGAGCTC 4

RESULT 6
 CG532825/c
 LOCUS CG532825 73 bp DNA linear GSS 01-OCT-2003
 DEFINITION OStr117228 Mus musculus 129Sv/Ev Mus musculus genomic clone
 OStr117228, Genomic survey sequence.
 ACCESSION CG532825
 VERSION CG532825.1 GI:37319397
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE 1 (bases 1 to 73)
 AUTHORS Zambronic,B.P., Abuin,A., Ramirez-Solis,R., Richer,L.J.,
 Pigott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
 Fiddle,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
 Key,B.W., Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D.,
 Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,
 Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,
 Zhu,Q., Person,C. and Sands,A.T.
 Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
 screen to identify potential targets for therapeutic intervention
 Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
 CONTACT: Zambrowicz BP
 OmniBank
 Lexicon Genetics Incorporated
 4000 Research Forest Drive, The Woodlands, TX 77381, USA
 Email: materials@lexgen.com
 Gene trap sequence tag generated by 3' RACE from mouse ES cells as
 described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):1608-11)

```

FEATURES
  source
    Class: Gene Trap.
    Location/Qualifiers
      1..73
        /organism="Mus musculus"
        /mol_type="genomic DNA"
        /strain="129SV/Ev"
        /db_xref="taxon:10090"
        /clone="OST117228"
        /cell_type="embryonic stem cell"
        /clone_lib="Mus musculus 129SV/Ev"

ORIGIN
  Query Match      56.4%; Score 15.8; DB 9; Length 91;
  Best Local Similarity 74.1%; Pred. No. 2.6e+04;
  Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 GTCGGAAGTGCATCCAAATTCAGGTTTC 28
Db 60 GTAATGNACTGCATGCACACTCAAGTTC 33

RESULT 7
LOCUS      AA814152          91 bp      mRNA      linear      EST 05-MAR-1998
DEFINITION O025E02.51 NCI_CGAP Kid5 Homo sapiens cDNA clone IMAGE1324730 3',
            mRNA sequence.
ACCESSION  AA814152
VERSION     AA814152.1 GI:2883748
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 91)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs@mail.nih.gov
            unknown library type
            Insert Length: 1415 Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 77.
FEATURES
  source
    1..91
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:1324730"
      /tissue_type="2 pooled tumors (clear cell type)"
      /lab_host="DH10B"
      /clone_lib="NCI CGAP Kids"
      /notes="Organ: kidney; Vector: pTTT3D-Pac (Pharmacia) with
      a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
      strand cDNA was primed with a Not I - oligo(dT) primer [5',
      AACTGGAAGATTCGGCGCGCAATATTTTGTGTTT 3'],
      double-stranded cDNA was ligated to Eco RI adaptors
      (Pharmacia), digested with Not I and cloned into the Not I
      and Eco RI sites of the modified pTTT3 vector. Library
      went through one round of normalization. Library
      constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN
  Query Match      56.4%; Score 15.8; DB 9; Length 91;
  Best Local Similarity 74.1%; Pred. No. 2.6e+04;
  Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 TCTGGAAGTGCATCCAAATTCAGGTTTC 28
Db 69 TCTGGTACTAGCTACAAATTCGGTTTC 43

FEATURES
  source
    Class: Gene Trap.
    Location/Qualifiers
      1..73
        /organism="Mus musculus"
        /mol_type="genomic DNA"
        /strain="129SV/Ev"
        /db_xref="taxon:10090"
        /clone="OST117228"
        /cell_type="embryonic stem cell"
        /clone_lib="Mus musculus 129SV/Ev"

ORIGIN
  Query Match      56.4%; Score 15.8; DB 9; Length 91;
  Best Local Similarity 74.1%; Pred. No. 2.6e+04;
  Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 GTCGGAAGTGCATCCAAATTCAGGTTTC 28
Db 60 GTAATGNACTGCATGCACACTCAAGTTC 33

RESULT 8
LOCUS      AZ761676          100 bp      DNA      linear      GSS 16-FEB-2001
DEFINITION clone UUGC1M0556D07 F, genomic survey sequence.
ACCESSION  AZ761676
VERSION     AZ761676.1 GI:12870848
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 100)
            Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D., Weiss,R.,
            Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
            Unpublished (2000)
            Contact: Robert B. Weiss
            University of Utah Genome Center
            Rm 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0556 row: D column: 07
            Seq primer: CGTTGTAACGACGCGCCAGT
            Class: plasmid ends
            High quality sequence stop: 100.
FEATURES
  source
    1..100
      /organism="Mus musculus"
      /mol_type="genomic DNA"
      /strain="C57BL/6J"
      /db_xref="taxon:10090"
      /clone="UUGC1M0556D07"
      /sex="Male"
      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
      /clone_lib="Mouse 10kb plasmid UUGC1M library"
      /note="Vector: PWD42nv; Purified genomic DNA from M.
      musculus C57BL/6J (male) was obtained from the Jackson
      Laboratory Mouse DNA Resource
      (http://www.jax.org/resources/documents/dnares/). The DNA
      was hydrodynamically sheared by repeated passage through a
      0.005 inch orifice at constant velocity. The sheared DNA
      was blunt end-repaired with T4 DNA polymerase and T4
      polynucleotide kinase. Adaptor oligonucleotides were
      ligated to the blunt ends in high molar excess. The
      adaptor DNA was purified and size-selected for a 9.5 to
      10.5 kb range using preparative agarose gel
      electrophoresis. Vector DNA was prepared from a derivative
      of pWD42 [gi|4732114|gb|AF129072.1], a copy-number
      inducible derivative of plasmid R1. The vector was ligated
      with adaptors complementary to the insert adaptors and
      purified. The sheared, adaptor mouse DNA was annealed to
      adaptor vector DNA, and transformed into
      chemically-competent E. coli XL10-Gold (Stratagene) cells
      and selected for ampicillin resistance."

ORIGIN
  Query Match      56.4%; Score 15.8; DB 28; Length 100;
  Best Local Similarity 74.1%; Pred. No. 2.7e+04;
  Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 TCTGGAAGTGCATCCAAATTCAGGTTTC 28
Db 25 TATGGAATAGCATCCAAATTAATGATC 51

```

```
RESULT 9
CD946266/c
LOCUS          CD946266          66 bp    mRNA    linear    EST 15-JUL-2003
DEFINITION    REL 29 Genetag1 Zea mays cDNA, mRNA sequence.
ACCESSION     CD946266
VERSION       CD946266.1 GI:32794030
KEYWORDS      EST.
SOURCE        Zea mays
ORGANISM      Zea mays
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
              clade; Panicoideae; Andropogoneae; Zea.
              1 (bases 1 to 66)
REFERENCE
AUTHORS       Genoplante.
TITLE         Genoplante, a major partnership french program in plant genomics
JOURNAL       Unpublished (2003)
COMMENT       Contact: Genoplante
              Genoplante
              93, rue Henri Rochefort 91025 EVRY CEDEX France
              Tel: 33 1 69 47 54 00
              Fax: 33 1 69 47 54 10
              This sequence has been generated in the framework of the french
              plant genomics programme 'Genoplante' (http://www.genoplante.com)
              and http://genoplante-info.infobiogen.fr.
FEATURES
source        Location/Qualifiers
              1..66
              /organism="Zea mays"
              /mol_type="mRNA"
              /cultivar="mixture"
              /db_xref="taxon:4577"
              /clone_lib="Genetag1"
ORIGIN
Query Match          55.7%; Score 15.6; DB 14; Length 66;
Best Local Similarity 81.8%; Pred. No. 2.9e+04;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy  3  CTGGAACCTGCATCCAAATTCAG 24
      |||||
Db   37  CAGAACCGCATCCACATTCAG 16

RESULT 10
CD946207/c
LOCUS          CD946207          67 bp    mRNA    linear    EST 15-JUL-2003
DEFINITION    REK_16 Genetag1 Zea mays cDNA, mRNA sequence.
ACCESSION     CD946207
VERSION       CD946207.1 GI:32793971
KEYWORDS      EST.
SOURCE        Zea mays
ORGANISM      Zea mays
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
              clade; Panicoideae; Andropogoneae; Zea.
              1 (bases 1 to 67)
REFERENCE
AUTHORS       Genoplante.
TITLE         Genoplante, a major partnership french program in plant genomics
JOURNAL       Unpublished (2003)
COMMENT       Contact: Genoplante
              Genoplante
              93, rue Henri Rochefort 91025 EVRY CEDEX France
              Tel: 33 1 69 47 54 00
              Fax: 33 1 69 47 54 10
              This sequence has been generated in the framework of the french
              plant genomics programme 'Genoplante' (http://www.genoplante.com)
              and http://genoplante-info.infobiogen.fr.
FEATURES
source        Location/Qualifiers
              1..67
              /organism="Zea mays"
              /mol_type="mRNA"
              /cultivar="mixture"
              /db_xref="taxon:4577"
              /clone_lib="Genetag1"
ORIGIN
Query Match          55.7%; Score 15.6; DB 14; Length 66;
Best Local Similarity 81.8%; Pred. No. 2.9e+04;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy  3  CTGGAACCTGCATCCAAATTCAG 24
      |||||
Db   37  CAGAACCGCATCCACATTCAG 16

RESULT 11
AW268320/c
LOCUS          AW268320          56 bp    mRNA    linear    EST 03-JAN-2000
DEFINITION    xr95f11.x1 NCI_CGAP_Lu26 Homo sapiens cDNA clone IMAGE:2767917 3',
              mRNA sequence.
ACCESSION     AW268320
VERSION       AW268320.1 GI:6655350
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              1 (bases 1 to 56)
REFERENCE
AUTHORS       NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
ORIGIN
Query Match          54.3%; Score 15.2; DB 9; Length 50;
Best Local Similarity 71.4%; Pred. No. 3.9e+04;
Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy  1  GTCTGGAACCTGCATCCAAATTCAGGTTTC 28
      |||||
Db   5  GCCTGCATCTGTATCCAGGTTCAGGTCC 32

RESULT 12
AW268320/c
LOCUS          AW268320          56 bp    mRNA    linear    EST 03-JAN-2000
DEFINITION    xr95f11.x1 NCI_CGAP_Lu26 Homo sapiens cDNA clone IMAGE:2767917 3',
              mRNA sequence.
ACCESSION     AW268320
VERSION       AW268320.1 GI:6655350
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              1 (bases 1 to 56)
REFERENCE
AUTHORS       NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
ORIGIN
Query Match          54.3%; Score 15.2; DB 9; Length 50;
Best Local Similarity 71.4%; Pred. No. 3.9e+04;
Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy  1  GTCTGGAACCTGCATCCAAATTCAGGTTTC 28
      |||||
Db   5  GCCTGCATCTGTATCCAGGTTCAGGTCC 32
```

```
ORIGIN
Query Match          55.7%; Score 15.6; DB 14; Length 67;
Best Local Similarity 81.8%; Pred. No. 2.9e+04;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy  3  CTGGAACCTGCATCCAAATTCAG 24
      |||||
Db   30  CAGAACCGCATCCACATTCAG 51

RESULT 11
AW268320/c
LOCUS          AW268320          50 bp    mRNA    linear    EST 30-AUG-2001
DEFINITION    KAT08774, mRNA sequence.
ACCESSION     AU105882
VERSION       AU105882.1 GI:13555403
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              1 (bases 1 to 50)
REFERENCE
AUTHORS       Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J.,
              Hata, H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K.,
              Sakaki, Y., Nakamura, Y., Suyama, A. and Sugano, S.
              Diverse transcriptional initiation revealed by fine, large-scale
              mapping of mRNA start sites
              EMBO Rep. 2 (5), 388-393 (2001)
              21270072
              PUBMED 11375929
              CONTACT: Yutaka Suzuki
              Department of Virology
              Institute of Medical Science, University of Tokyo
              4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
              Email: ysuzuki@ims.u-tokyo.ac.jp
              Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and
              Sugano, S. Construction and characterization of a full
              length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
              149-156 (1997).
              Location/Qualifiers
              1..50
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone_lib="KAT08774"
              /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN
Query Match          54.3%; Score 15.2; DB 9; Length 50;
Best Local Similarity 71.4%; Pred. No. 3.9e+04;
Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy  1  GTCTGGAACCTGCATCCAAATTCAGGTTTC 28
      |||||
Db   5  GCCTGCATCTGTATCCAGGTTCAGGTCC 32

RESULT 12
AW268320/c
LOCUS          AW268320          56 bp    mRNA    linear    EST 03-JAN-2000
DEFINITION    xr95f11.x1 NCI_CGAP_Lu26 Homo sapiens cDNA clone IMAGE:2767917 3',
              mRNA sequence.
ACCESSION     AW268320
VERSION       AW268320.1 GI:6655350
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              1 (bases 1 to 56)
REFERENCE
AUTHORS       NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
ORIGIN
Query Match          54.3%; Score 15.2; DB 9; Length 50;
Best Local Similarity 71.4%; Pred. No. 3.9e+04;
Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy  1  GTCTGGAACCTGCATCCAAATTCAGGTTTC 28
      |||||
Db   5  GCCTGCATCTGTATCCAGGTTCAGGTCC 32
```

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: I.M.A.G.E. Consortium, LLNL
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bic.llnl.gov/bicrp/image/image.html

Possible reversed clone: polyT not found

Seq primer: -40UP from Gibco.

FEATURES

source
Location/Qualifiers

1..56
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2767917"
/tissue_type="invasive adenocarcinoma"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu26"
/note="Organ: lung; Vector: pAMP1; mRNA made from lung
adenocarcinoma tissue, cDNA made by oligo-dT priming.
Directionally cloned. Size-selected on agarose gel,
average insert size 500 bp. Primary library,
non-amplified."

ORIGIN

Query Match 54.3%; Score 15.2; DB 10; Length 56;
Best Local Similarity 71.4%; Pred. No. 4.1e+04;
Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 GTCTGGAAGTCATCCCAAAATTCAGGTC 28
||||| ||| ||||| |||||
DB 25 GTCTGCCCCCTCATTAATAATTCAGGTC 52

RESULT 13
BJ041364/C
LOCUS BJ041364 59 bp mRNA linear EST 29-SEP-2003
DEFINITION BJ041364 NIBB Mochii normalized xenopus neurula library Xenopus
laevis cDNA clone XL008e05 5', mRNA sequence.
ACCESSION BJ041364
VERSION BJ041364.1 GI:17412190
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)

ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 59)
REFERENCE Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-i, T. and
Kohara, Y.
Expressed genes in X. laevis embryo
Unpublished (2001)
TITLE Contact: Tadasi Shin-i
JOURNAL Center For Genetic Resource Information
COMMENT National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp
The information of this clone is available through the following
URL:
http://xenopus.nibb.ac.jp.

FEATURES source
Location/Qualifiers
1..59
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"

/clone="XL008e05"
/tissue_type="whole embryo"
/dev_stage="stage 15"
/clone_lib="NIBB Mochii normalized Xenopus neurula
library"

ORIGIN

Query Match 54.3%; Score 15.2; DB 12; Length 59;
Best Local Similarity 85.0%; Pred. No. 4.1e+04;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 TCTGGAAGTCATCCCAAAAT 21
||||| ||||| |||||
DB 23 TCTGATCTGCATCCAAAT 4

RESULT 14

AI174615
LOCUS AI174615 61 bp mRNA linear EST 07-OCT-1998
DEFINITION an47d11.s1 Gessler Wilms tumor Homo sapiens cDNA clone
IMAGE:1701813 3', similar to SW:ATP6_TRYBB P24499 ATP SYNTHASE A
CHAIN;; mRNA sequence.
ACCESSION AI174615
VERSION AI174615.1 GI:3721468
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 61)
REFERENCE Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Scheilberg, K., Steptoe, M., Tan, F.,
Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)
TITLE Contact: Wilson RK
JOURNAL Washington University School of Medicine
COMMENT 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..61
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1701813"
/sex="pooled (6)"
/lab_host="DH10B"
/clone_lib="Gessler Wilms tumor"
/note="Vector: pSPOR11; Site: 1; SalI; Site 2: NotI; RNA
was prepared from a pool of 6 anonymous Wilms' tumor RNAs.
RNA was prepared by acid-phenol, followed by one round of
oligo dT selection. cDNA library preparation was with
the BRL/Life Tech. Superscript Plasmid system. An
oligo-dT NotI primer for first strand synthesis generated
9cggccggccctn at the 3' end of the clones. A 5' SalI
adaptor was used with sequence 5'-gtcggaccacgctcg-3'.
Resulting cDNAs were size selected (average size 2 Kb),
NotI digested, and ligated into NotI/SalI-cut pSPORT1.
Library was constructed by Dr. Manfred Gessler."

ORIGIN

Query Match 54.3%; Score 15.2; DB 9; Length 61;
Best Local Similarity 85.0%; Pred. No. 4.1e+04;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 AACTGCATCCAAATTCAGGT 26
|||||
Db 17 AACTGAATCCCAAGTCAGGT 36
|||||

RESULT 15
AA671040/c
LOCUS
DEFINITION
AA671040 76 bp mRNA linear EST 25-NOV-1997
vis1f01.r1 Beddington mouse embryonic region Mus musculus cDNA
clone IMAGE:907321 5' similar to TR:G192989 G192989 ELONGATION
FACTOR 2 ; mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)

REFERENCE
AUTHORS
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 76)
Marra.M., Hillier.L., Allen.M., Bowles.M., Dietrich.N., Dubuque.T.,
Geisel.S., Kucaba.T., Lacy.M., Le.M., Martin.J., Morris.M.,
Schellenberg.K., Steptoe.K., Tan.F., Underwood.K., Moore.B.,
Theising.B., Wylie.F., Lennon.G., Soares.B., Wilson.R. and
Waterston.R.

TITLE
JOURNAL
COMMENT
The WashU-HHMI Mouse EST Project
Unpublished (1996)
Contact: Marta M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG:527985

Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 1.

FEATURES
Location/Qualifiers
1..76
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6 x DBA"
/db_xref="taxon:10090"
/clone="IMAGE:907321"
/sex="pooled"
/tissue_type="embryo"
/dev_stages="7.5dpc"
/lab_host="DH12S"
/clone_lib="Beddington mouse embryonic region"
/note="Organ: whole embryo; Vector: pCMV-SPORT; Site 1:
Salt; Site 2: NotI; Cloned unidirectionally. Primer:
Oligo dT. Gastrulating embryos were collected at 7.5dpc
from C57BL/6 x DBA matings, excluding embryos that had
developed head folds and all extraembryonic tissues.
Average insert size: 1.3 kb (range: 0.5 - 3.0 kb).
Referenced in Development 121, 2479-2489 (1995)";

ORIGIN
Query Match 54.3%; Score 15.2; DB 9; Length 76;
Best Local Similarity 85.0%; Pred.No.4.4e+04;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 GAATGCATCCAAATTCAGG 25
|||||
Db 32 GAACCGCATCAAAATGCAGG 13
|||||

Search completed: September 1, 2004, 23:38:32
Job time : 1769 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2004, 17:24:51 ; Search time 233 Seconds
(without alignments)
510.513 Million cell updates/sec

Title: US-10-028-415-2
Perfect score: 28
Sequence: 1 gctggaactgcatccaaattcagggtc 28

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 337363 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

- 1: geneseqn1980s.*
- 2: geneseqn1990s.*
- 3: geneseqn2000s.*
- 4: geneseqn2001as.*
- 5: geneseqn2001bs.*
- 6: geneseqn2002as.*
- 7: geneseqn2003as.*
- 8: geneseqn2003bs.*
- 9: geneseqn2003cs.*
- 10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	100.0	28	2 AAV20897	AAV20897 CD95 regu
C 2	28	100.0	28	2 AAV23143	AAV23143 CD95 regu
C 3	28	100.0	28	6 ABK88709	ABK88709 Human CD9
4	28	100.0	28	6 ABK88700	ABK88700 Transcrip
5	28	100.0	50	2 AAV23172	AAV23172 CD95 enha
6	28	100.0	1608	2 AAT34162	AAT34162 Fas promo
7	28	100.0	2165	6 ABN79677	ABN79677 Sequence
8	28	100.0	28118	9 ADE43798	ADE43798 Polymorph
9	28	100.0	28118	9 ADE43797	ADE43797 Human TNF
10	28	100.0	45121	8 ADA02744	ADA02744 Human TNF
11	28	100.0	45121	9 ADE72482	ADE72482 Human TNF
12	28	100.0	45121	9 ADC85224	ADC85224 Human Tnf
C 13	20	71.4	9193	6 AAD46346	AAV23143 CD95 regu
C 14	19	67.9	1608	6 ABL34419	ABL34419 Human imm
C 15	19	67.9	108359	8 ADA13316	ADA13316 Human fri
C 16	19	67.9	160771	6 ABQ88179	ABQ88179 Human ost
C 17	18.8	67.1	1424	6 ABZ33746	ABZ33746 Human TRI
C 18	18.6	66.4	460	6 ABN21606	ABN21606 Human ORF
C 19	18.6	66.4	508	8 ACH36880	ACH36880 Human end
C 20	18.6	66.4	525	4 AHO9866	AHO9866 Human CDN
C 21	18.6	66.4	1487	4 AAH13934	AAH13934 Human CDN
C 22	18.6	66.4	2287	5 ABA14736	ABA14736 Human ner
C 23	18.6	66.4	2287	5 ABA14735	ABA14735 Human ner

ALIGNMENTS

RESULT 1

AAV20897

ID AAV20897 standard; DNA; 28 BP.

XX AC AAV20897;

XX AC AAV20897;

DT 23-JUL-1998 (first entry)

XX DE CD95 regulatory region.

XX XX

XX CD95; regulatory region; silencer; enhancer; apoptosis; inhibitor;
cancer; viral infection; neurodegeneration; autoimmune disease;
gene therapy; transcription factor; ss.

XX OS Homo sapiens.

XX OS Homo sapiens.

XX WO9808965-A2.

XX 05-MAR-1998.

XX 29-AUG-1997; 97WO-NZ000107.

XX 30-AUG-1996; 96US-00713557.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX Watson JD, Rudert F;

XX WPI; 1998-179445/16.

XX New regulatory regions from the CD95 gene and transcription factors that

XX interact with them - for control of apoptosis, e.g. in treatment of

XX cancer, viral infection, neurodegeneration and autoimmune disease.

XX Claim 1; Page 36; 60pp; English.

XX This sequence represents a regulatory region from the CD95 gene, and is a

XX nucleic acid of the invention. The regulatory regions (silencers or

XX enhancers) are involved in apoptosis, i.e. inhibition of CD95 expression

XX will inhibit apoptosis. Regulation of apoptosis is useful in treatment of

XX cancer, (retroviral infection, neurodegeneration and autoimmune disease,

XX e.g. by gene therapy for expressing transcription factors or expression

XX of antisense sequences to inhibit transcription factor production. The

XX regulatory nucleic acids and transcription factors are also useful for

XX studying regulation of CD95 in vitro or in vivo, to screen for modulators

XX and as probes to isolate related genes

AAZ58978 Human cyt
AAS45004 CDNA enco
ABH35087 Human end
AAZ34186 Human EST
AAZ78546 Human EST
AAZ93396 Aortic en
ACD42719 Novel hum
ACA63754 Novel hum
ACA71918 Human PRO
ABX92558 Human PRO
ACA66299 Human sec
ADA24886 Novel hum
ADA29300 Novel hum
ADA12547 Human sec
ADC29315 Novel hum
ADH73853 Human PRO
ADH76569 Human PRO
ADC43995 Human EST
ADC61755 Human EST
ADC63719 Human EST
ADC66819 Human EST
ADC68943 Human EST

C 24 18.6 66.4 3191 3 AAZ58978
C 25 18.6 66.4 3363 5 AAS45004
C 26 18.4 65.7 501 2 ABH35087
C 27 18.4 65.7 509 8 AAZ34186
C 28 18.4 65.7 509 3 AAZ78546
C 29 18.4 65.7 509 5 AAZ93396
C 30 18.4 65.7 509 7 ACD42719
C 31 18.4 65.7 509 7 ACA63754
C 32 18.4 65.7 509 7 ACA71918
C 33 18.4 65.7 509 7 ABX92558
C 34 18.4 65.7 509 7 ACA66299
C 35 18.4 65.7 509 8 ADA24886
C 36 18.4 65.7 509 8 ADA29300
C 37 18.4 65.7 509 8 ADA12547
C 38 18.4 65.7 509 8 ADC29315
C 39 18.4 65.7 509 9 ADH73853
C 40 18.4 65.7 509 9 ADH76569
C 41 18.4 65.7 509 9 ADC43995
C 42 18.4 65.7 509 9 ADC61755
C 43 18.4 65.7 509 9 ADC63719
C 44 18.4 65.7 509 9 ADC66819
C 45 18.4 65.7 509 9 ADC68943


```
AC ABK88700;
XX
XX
DT 07-OCT-2002 (first entry)
XX
XX
DE Transcription silencer region (S1) of human CD95 gene.
XX
XX Human; apoptotic cell death; proteinaceous transcription factor;
XX regulation of gene transcription; apoptosis; p53; CD95; TRA;
XX transcriptional regulator of apoptosis; Y-box family; YB-1; cancer;
XX tumour cell; embryonic cell; nervous system; intracellular pathogen;
XX DNA-damaging agent; retroviral infection; neurodegenerative disorder;
XX immune system dysfunction; anti-tumour; cytostatic; S1; hCD95;
XX transcription silencer region; ds.
XX
XX Homo sapiens.
XX
XX WO200244363-A1.
XX
XX
PD 06-JUN-2002.
XX
XX
XX 28-NOV-2001; 2001WO-NZ000287.
XX
XX 28-NOV-2000; 2000US-00724809.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX
XX Lasham A, Watson JD;
XX
XX WPI; 2002-557540/59.
XX
XX Modulating p53-mediated apoptotic cell death in a population of cells, by
XX modulating the amount of a transcriptional regulator of apoptosis
XX available to bind to a target polynucleotide in the cells.
XX
XX Example 1; Page 53; 62pp; English.
XX
XX The present invention relates to methods for modulating apoptotic cell
XX death using proteinaceous transcription factors that regulate the
XX transcription of genes encoding proteins involved in apoptosis (e.g. CD95
XX and p53). The methods involve modulating the amount of a transcriptional
XX regulator of apoptosis (TRA) available to bind to a target polynucleotide
XX in the cells, where TRA is a member of the Y-box nucleic acid binding
XX family of polypeptides (e.g. YB-1). The methods of the invention are
XX useful for modulating apoptotic cell death in a population of cells,
XX where the cells are selected from tumour cells, cells of the immune
XX system, embryonic cells, cells of the nervous system, or cells infected
XX with intracellular pathogens. The methods are also useful for increasing
XX the sensitivity of tumour cells to a DNA-damaging agent, and for
XX increasing sensitivity to apoptosis in a population of cells harbouring
XX intracellular pathogens. The methods are useful for screening an
XX apoptosis modulatory agent that modulates the binding of TRA. The methods
XX for regulating apoptosis can be used therapeutically and prophylactically
XX for various disorders such as cancer, viral and retroviral infections,
XX neurodegenerative disorders, and immune system dysfunction. The present
XX sequence represents the transcription silencer region (S1) of the human
XX CD95 (hCD95) gene
XX
XX Sequence 28 BP; 7 A; 7 C; 6 G; 8 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 28; DB 6; Length 28;
XX Best Local Similarity 100.0%; Pred. No. 0.0064;
XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GTCGTGAAGTGCATCCAAATTCAGGTTTC 28
XX
XX Db 1 GTCGTGAAGTGCATCCAAATTCAGGTTTC 28
XX
XX
XX RESULT 5
XX AAV23172
XX ID AAV23172 standard; DNA; 50 BP.
XX
XX
XX AAV23172;
XX
```

```
XX
XX 23-JUL-1998 (first entry)
XX
XX CD95 enhancer binding site.
XX
XX CD95; regulatory region; silencer; enhancer; apoptosis; inhibitor;
XX cancer; viral infection; neurodegeneration; autoimmune disease;
XX gene therapy; transcription factor; ss.
XX
XX Homo sapiens.
XX
XX WO9808965-A2.
XX
XX 05-MAR-1998.
XX
XX 29-AUG-1997; 97WO-NZ000107.
XX
XX 30-AUG-1996; 96US-00713557.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX
XX Watson JD, Rudert F;
XX
XX WPI; 1998-179445/16.
XX
XX New regulatory regions from the CD95 gene and transcription factors that
XX interact with them - for control of apoptosis, e.g. in treatment of
XX cancer, viral infection, neurodegeneration and autoimmune disease.
XX
XX Disclosure; Page 43; 60pp; English.
XX
XX This sequence represents a regulatory region from the CD95 gene, and is a
XX nucleic acid of the invention. The regulatory regions (silencers or
XX enhancers) are involved in apoptosis, i.e. inhibition of CD95 expression
XX will inhibit apoptosis. Regulation of apoptosis is useful in treatment of
XX cancer, (retro)viral infection, neurodegeneration and autoimmune disease,
XX e.g. by gene therapy for expressing transcription factors or expression
XX of antisense sequences to inhibit transcription factor production. The
XX regulatory nucleic acids and transcription factors are also useful for
XX studying regulation of CD95 in vitro or in vivo, to screen for modulators
XX and as probes to isolate related genes
XX
XX Sequence 50 BP; 16 A; 10 C; 10 G; 14 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 28; DB 2; Length 50;
XX Best Local Similarity 100.0%; Pred. No. 0.007;
XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GTCGTGAAGTGCATCCAAATTCAGGTTTC 28
XX
XX Db 1 GTCGTGAAGTGCATCCAAATTCAGGTTTC 28
XX
XX
XX RESULT 6
XX AAT34162
XX ID AAT34162 standard; DNA; 1608 BP.
XX
XX AC AAT34162;
XX
XX 22-OCT-1996 (first entry)
XX
XX DE Fas promoter region.
XX
XX Fas gene promoter; apoptosis; ageing; autoimmune disease;
XX T-cell senescence; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX promoter 1..1074
XX /*tag= a
XX /note= "Fas gene promoter region"
XX protein_bind 147..151
```

FT /tag= b
 FT /function= "GFI transcription factor binding site"
 FT /note= "Claim 7"
 FT protein_bind 168..174
 FT /tag= c
 FT /function= "EBF20 transcription factor binding site"
 FT protein_bind 272..276
 FT /tag= d
 FT /function= "Myb transcription factor binding site"
 FT protein_bind 349..353
 FT /note= "Claim 9"
 FT /tag= e
 FT /function= "NF-Y transcription factor binding site"
 FT protein_bind 521..525
 FT /tag= f
 FT /function= "GF-1 transcription factor binding site"
 FT protein_bind 604..609
 FT /tag= g
 FT /function= "NF-Y transcription factor binding site"
 FT protein_bind 621..626
 FT /note= "Claim 8"
 FT /tag= h
 FT /function= "AP-1 transcription factor binding site"
 FT protein_bind 1037..1043
 FT /tag= i
 FT /function= "CP2 transcription factor binding site"
 FT exon 1075..1476
 FT /tag= j
 FT /codon_start= 1479..1469
 FT /product= "Fas protein leader"
 FT intron 1497..1608
 FT /tag= j
 FT /note= "5' end of intron 1 (full length approx. 14 kb)"
 FT XX
 FT WO9622370-A1.
 FT XX
 FT 25-JUL-1996.
 FT PD
 FT 19-JAN-1996; 96WO-US000606.
 FT PF
 FT 20-JAN-1995; 95US-00377522.
 FT PR
 FT (UABR-) UAB RES FOUND.
 FT PA
 FT Mountz JD, Liu C, Cheng J, Koopman WJ, Zhou T;
 FT PI
 FT WPI; 1996-354527/35.
 FT DR P-PSDB; AAR39471.
 FT XX
 FT Human Fas gene promoter region - used for heterologous protein expression
 FT PT and for developing products for treating Fas-mediated apoptosis
 FT PT disorders.
 FT XX
 FT Claim 11; Fig 2; 123pp; English.
 FT PS
 FT A novel DNA segment (AAR34162) has an isolated sequence region defined as
 FT CC the Fas gene promoter region. This includes a number of transcription
 FT CC factor binding sites. A coding sequence for the N-terminal portion
 FT CC (AAR39471) of the Fas protein leader peptide is also included. The DNA
 FT CC segment was isolated from a human placental DNA library using a 32p-
 FT CC labeled segment of human Fas cDNA corresponding to nt 23-346. It can be
 FT CC combined with a structural gene so that the gene is under the
 FT CC transcriptional control of the transcription factor binding sites. The
 FT CC promoter region can be used to regulate Fas gene expression. e.g. in
 FT CC tumour or immune cells, as a means of treating Fas-mediated apoptosis
 FT CC disorders such as malignancies and autoimmune diseases
 FT XX
 FT Sequence 1608 BP; 398 A; 421 C; 423 G; 366 T; 0 U; 0 Other;

Query Match 100.0%; Score 28; DB 2; Length 1608;
 Best Local Similarity 100.0%; Pred. No. 0.012;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTCGGAAGTGCATCCAAATTCAGGTTTC 28
 Db 433 GTCGGAAGTGCATCCAAATTCAGGTTTC 460
 RESULT 7
 ABN79677
 ID ABN79677 standard; DNA; 2165 BP.
 XX
 AC ABN79677;
 XX
 DT 29-JUL-2002 (first entry)
 XX
 DE Sequence #1 used to generate target oligonucleotides.
 XX
 KW Human; immunosuppressive; antiinflammatory; hepatotropic; cytostatic;
 KW vasotropic; hepatitis; cancer; allograft rejection; ds; Fas.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1782..1813
 FT /tag= a
 FT /product= "peptide encoded by sequence used create target
 FT oligonucleotides"
 XX
 PN US2002004490-A1.
 XX
 PD 10-JAN-2002.
 XX
 PF 09-MAR-2001; 2001US-00802669.
 PR 12-APR-1999; 99US-00290640.
 PR 18-SEP-2000; 2000US-00665615.
 XX
 PA (DEAN/) DEAN N M.
 PA (MARC/) MARCUSSEN E G.
 PA (WYAT/) WYATT J.
 PA (ZHAN/) ZHANG H.
 XX
 PI Dean NM, Marcussen EG, Wyatt J, Zhang H;
 XX
 DR WPI; 2002-204886/26.
 DR P-PSDB; ABP35566.
 XX
 PT Novel antisense compound targeted to nucleic acid encoding Fas, Fas
 PT ligand or Fas associated protein-1 is useful for inhibiting expression of
 PT Fas, Fas ligand, or Fas-1 in cells or tissues, and for treating
 PT hepatitis.
 XX
 PS Example 18; Page 62-63; 84pp; English.
 XX
 CC This invention relates to an antisense compound encoding Fas, Fas ligand,
 CC or Fas associated protein-1 (Fap-1). The inhibition of Fas mediated
 CC signalling is thought to be immunosuppressive, antiinflammatory,
 CC hepatotropic, cytostatic and vasotropic. Antisense oligonucleotides were
 CC designed to target human Fas. Oligonucleotides were synthesised as
 CC chimeric oligonucleotides and are useful for treating an animal having an
 CC autoimmune or inflammatory disease e.g., hepatitis, cancer, a condition
 CC associated with apoptosis, allograft rejection, or ischemia reperfusion
 CC injury. Optionally, the above mentioned conditions are prevented by
 CC contacting the allograft with the antisense oligonucleotide. The
 CC oligonucleotides are used in diagnostics, therapeutics, prophylaxis and
 CC as research reagents and in kits. The oligonucleotides are also useful
 CC for research purposes. The present nucleotide sequence is related to
 CC human Fas
 XX
 SQ Sequence 2165 BP; 508 A; 555 C; 555 G; 547 T; 0 U; 0 Other;

```
Query Match      100.0%; Score 28; DB 6; Length 2165;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGGAAGTGCATCCAAATTCAGGTTTC 28
Db 747 GTCGGAAGTGCATCCAAATTCAGGTTTC 774

RESULT 8
ADE43798
ID ADE43798 standard; DNA; 28118 BP.
XX
XX ADE43798;
XX
XX 29-JAN-2004 (first entry)
XX
XX Polymorphic human TNFRSF6 genomic sequence, SEQ ID 403.
XX
XX Neurodegenerative disease; uPA; SNGG; IDE; KNSLI; LIPA; TNFRSF6;
XX Alzheimer's disease; neuroprotective; nootropic; gene therapy;
XX Chromosome 10; gene; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH misc_feature 199
FT /tag= a
FT /note= "There is a variation at this position"
FT 213
FT /tag= b
FT /note= "There is a variation at this position"
FT 843
FT /tag= c
FT /note= "There is a variation at this position"
FT 1530
FT /tag= d
FT /note= "There is a variation at this position"
FT 1550
FT /tag= e
FT /note= "There is a variation at this position"
FT 2967
FT /tag= f
FT /note= "There is a variation at this position"
FT 3103
FT /tag= g
FT /note= "There is a variation at this position"
FT 5335
FT /tag= h
FT /note= "There is a variation at this position"
FT 5345
FT /tag= i
FT /note= "There is a variation at this position"
FT 6074
FT /tag= j
FT /note= "There is a variation at this position"
FT 9374
FT /tag= k
FT /note= "There is a variation at this position"
FT 9907
FT /tag= l
FT /note= "There is a variation at this position"
FT 9936
FT /tag= m
FT /note= "There is a variation at this position"
FT 10937
FT /tag= n
FT /note= "There is a variation at this position"
FT 11200
FT /tag= o
FT /note= "There is a variation at this position"
FT 11279
FT /tag= p
```

```
FT misc_feature /note= "There is a variation at this position"
FT 11359
FT /tag= q
FT /note= "There is a variation at this position"
FT 11503
FT /tag= r
FT /note= "There is a variation at this position"
FT 11511
FT /tag= s
FT /note= "There is a variation at this position"
FT 11587
FT /tag= t
FT /note= "There is a variation at this position"
FT 11694
FT /tag= u
FT /note= "There is a variation at this position"
FT 11905
FT /tag= v
FT /note= "There is a variation at this position"
FT 12193
FT /tag= w
FT /note= "There is a variation at this position"
FT 12208
FT /tag= x
FT /note= "There is a variation at this position"
FT 12238
FT /tag= y
FT /note= "There is a variation at this position"
FT 14525
FT /tag= z
FT /note= "There is a variation at this position"
FT 14714
FT /tag= aa
FT /note= "There is a variation at this position"
FT 18511
FT /tag= ab
FT /note= "There is a variation at this position"
FT 18567
FT /tag= ac
FT /note= "There is a variation at this position"
FT 18982
FT /tag= ad
FT /note= "There is a variation at this position"
FT 19069
FT /tag= ae
FT /note= "There is a variation at this position"
FT 20412
FT /tag= af
FT /note= "There is a variation at this position"
FT 20552
FT /tag= ag
FT /note= "There is a variation at this position"
FT 20640
FT /tag= ah
FT /note= "There is a variation at this position"
FT 21585
FT /tag= ai
FT /note= "There is a variation at this position"
FT 22439
FT /tag= aj
FT /note= "There is a variation at this position"
FT 23199
FT /tag= ak
FT /note= "There is a variation at this position"
FT 23416
FT /tag= al
FT /note= "There is a variation at this position"
FT 24890
FT /tag= am
FT /note= "There is a variation at this position"
FT 26359
FT /tag= an
FT /note= "There is a variation at this position"
```



```

XX PR 26-DEC-2001; 2001US-00035832.
XX PA (SAGR-) SAGRES DISCOVERY.
XX PI Morris DW;
XX DR WPI; 2003-587068/55.
XX PT New recombinant nucleic acid encoding carcinoma associated protein,
XX useful for preparing compositions for treating carcinomas.
XX PS Claim 1; SEQ ID NO 1262; 245pp; English.
XX CC The invention relates to recombinant carcinoma associated (CA) nucleic
XX acid sequences from mouse and human (ADA01482-ADA03094), and to
XX recombinant carcinoma associated proteins (CAP) encoded by them. The
XX invention also encompasses expression vectors and host cells comprising a
XX CA nucleic acid, a polypeptide (especially an antibody) that specifically
XX binds to the protein, and a biochip comprising CA nucleic acid or
XX fragments thereof. The sequences of the invention were identified using
XX onco-genic retroviruses, which insert into the genome of the host organism
XX at random. Many of these do not carry transduced host oncogenes or
XX pathogenic trans-acting viral genes, meaning that cancer incidence is a
XX direct consequence of the effects of proviral integration into host
XX proto-oncogenes. The CA nucleic acid sequences can be used to diagnose
XX carcinoma (especially breast cancer, prostate cancer, lymphoma or
XX leukaemia) or a propensity to carcinoma by determination of the sequence
XX of a CA gene, or by determination of CA gene expression in particular
XX tissues. CA nucleic acids, proteins and antibodies are also useful as
XX therapeutic agents and in screening and evaluating drug candidates. The
XX present sequence represents a specifically claimed human CA nucleic acid
XX sequence of the invention. Note: The complete sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 45121 BP; 13226 A; 8836 C; 9010 G; 14049 T; 0 U; 0 Other;

Query Match 100.0%; Score 28; DB 8; Length 45121;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGTGGAAGTCATCCAAATTCAGGTTTC 28
DB 9186 GTCGTGGAAGTCATCCAAATTCAGGTTTC 9213

RESULT 11
ADB72482
ID ADB72482 standard; DNA; 45121 BP.
XX AC ADB72482;
XX DT 04-DEC-2003 (first entry)
XX DE Human TNFRSF6 gene.
XX KW human; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
XX cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX OS Homo sapiens.
XX PN WO2003008583-A2.
XX PD 30-JAN-2003.
XX PF 26-DEC-2001; 2001WO-US051291.
XX PR 02-MAR-2001; 2001US-00798586.
XX PR 23-OCT-2001; 2001US-00004113.
XX PR 08-NOV-2001; 2001US-00052482.
XX PR 30-NOV-2001; 2001US-00997722.

```

```

PR 20-DEC-2001; 2001US-00034650.
XX PA (SAGR-) SAGRES DISCOVERY.
XX PI Morris DW, Engelhard EK;
XX DR WPI; 2003-239337/23.
XX PT New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
XX cancers, neoplasm, adenocarcinoma, or sarcomas.
XX PS Claim 1; SEQ ID NO 310; 2304pp; English.
XX CC The invention relates to a novel recombinant nucleic acid comprising a
XX nucleotide sequence selected from any of the 660 sequences fully defined
XX in the specification. A polynucleotide of the invention has cytostatic
XX activity, and may have a use in gene therapy, or in a vaccine. The
XX recombinant nucleic acids and polypeptides are useful for treating
XX carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
XX sarcomas. The present sequence represents a human gene of the invention.
XX SQ Sequence 45121 BP; 13226 A; 8836 C; 9010 G; 14049 T; 0 U; 0 Other;

Query Match 100.0%; Score 28; DB 9; Length 45121;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGTGGAAGTCATCCAAATTCAGGTTTC 28
DB 9186 GTCGTGGAAGTCATCCAAATTCAGGTTTC 9213

RESULT 12
ADC85224
ID ADC85224 standard; DNA; 45121 BP.
XX AC ADC85224;
XX DT 01-JAN-2004 (first entry)
XX DE Human TNFRSF6 genomic sequence.
XX KW Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;
XX secreted; transmembrane; intracellular; ds.
XX OS Homo sapiens.
XX PN WO2003045230-A2.
XX PD 05-JUN-2003.
XX PF 02-DEC-2002; 2002WO-US038582.
XX PR 30-NOV-2001; 2001US-00997722.
XX PA (SAGR-) SAGRES DISCOVERY.
XX PI Morris DW, Engelhard EK;
XX DR WPI; 2003-513603/48.
XX PT New recombinant nucleic acid comprising a nucleotide sequence of any of
XX the carcinoma-associated (CA) genes, useful for screening for drug
XX candidates for diagnosing or treating carcinomas.
XX PS Claim 1; SEQ ID NO 10; 983pp; English.
XX CC The invention relates to a recombinant nucleic acid comprising a
XX nucleotide sequence selected from any of the fully defined carcinoma-
XX associated (CA) genes from the 50 tables given in the specification. The
XX CA proteins are secreted, transmembrane or intracellular proteins. The
XX recombinant nucleic acids are useful for screening for drug candidates
XX for diagnosing or treating carcinomas. Sequences given in ADC85215-

```

CC ADC85514 represent CA genes of the invention.

XX SQ Sequence 45121 BP; 13226 A; 8836 C; 9010 G; 14049 T; 0 U; 0 Other;

Query Match 100.0%; Score 28; DB 9; Length 45121;

Best Local Similarity 100.0%; Pred. No. 0.021;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGGAAGTGCATCCAAATTCAGGTTTC 28

DB 9186 GTCGGAAGTGCATCCAAATTCAGGTTTC 9213

RESULT 13

AD46346/c

ID AAD46346 standard; DNA; 9193 BP.

XX AC AAD46346;

XX DT 27-JAN-2003 (first entry)

XX DE Human nuclear receptor L67 gene.

XX Human; nuclear receptor; L67; metabolic disorder; hormonal dysfunction;

KW immunological; cellular defence mechanism; immunomodulator; cytostatic;

KW arteriosclerosis; cell proliferation; neurosystemic disease; therapy;

KW cell aberration; cholesterol; differentiation; receptor; gene; ds.

XX OS Homo sapiens.

XX PN WO200270697-A1.

XX PD 12-SEP-2002.

XX PF 21-FEB-2002; 2002WO-EP001864.

XX PR 02-MAR-2001; 2001EP-00105098.

XX PA (LION-) LION BIOSCIENCE AG.

XX PI Casari G, Jackson D, Suckow J;

XX DR WPI; 2002-723265/78.

XX PT New nuclear receptor (L67) nucleic acids and proteins, useful for

PT developing and identifying compounds for treating disorders of cell

PT metabolism, homeostasis, proliferation and differentiation.

XX PS Claim 1; Page 52-57; 67pp; English.

XX CC The invention relates to nuclear receptor (L67) nucleic acids, proteins

CC and methods of use of L67 receptor. L67 is useful in identifying agents

CC that inhibit its cellular function. The method is useful for inhibiting

CC the cellular function of L67. The invention is useful for developing and

CC identifying compounds for treating metabolic disorders, hormonal

CC dysfunctions, immunological indications, neurosystemic diseases, high

CC cholesterol, or arteriosclerosis. L67 DNA is useful for making vectors

CC and for transforming cell that can be used for producing L67 protein. It

CC is also useful as scientific research tools for developing nucleic acid

CC probes for determining L67 expression levels, e.g. to identify disease or

CC abnormal states. It can also be used for developing analytical tools such

CC as antisense oligonucleotides for selectively inhibiting expression of

CC the L67 gene to determine physiological responses. L67 is useful for

CC screening for L67 antagonist and agonist activity for controlling

CC cellular functions such as cell proliferation, differentiation,

CC aberrations or cellular defence mechanisms. The present sequence is human

CC L67 gene

XX SQ Sequence 9193 BP; 2535 A; 2255 C; 2013 G; 2390 T; 0 U; 0 Other;

Query Match 71.4%; Score 20; DB 6; Length 9193;

Best Local Similarity 82.1%; Pred. No. 70;

Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GTCGGAAGTGCATCCAAATTCAGGTTTC 28

DB 9186 GTCGGAAGTGCATCCAAATTCAGGTTTC 9213

RESULT 14

ABL34419/c

ID ABL34419 standard; DNA; 1608 BP.

XX AC ABL34419;

XX DT 26-MAR-2002 (first entry)

XX DE Human immune system associated gene SEQ ID NO: 2392.

XX Human; immune system disease; cytosine methylation; antiasthmatic;

KW antiarteriosclerotic; antianaemic; cytostatic; neutropenic;

KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;

KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;

KW antineoplastic; cancer; eye disease; arteriosclerosis; anaemia;

KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;

KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;

ds.

XX OS Homo sapiens.

XX PN WO200200928-A2.

XX PD 03-JAN-2002.

XX PF 02-JUL-2001; 2001WO-EP007537.

XX PR 30-JUN-2000; 2000DE-01032529.

XX PR 01-SEP-2000; 2000DE-01043826.

XX PA (EPIC-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K;

XX DR WPI; 2002-130909/17.

XX PT Nucleic acid comprising fragment of chemically modified gene, useful for

PT diagnosis and treatment of diseases associated with abnormal cytosine

PT methylation.

XX PS Claim 1; SEQ ID NO 2392; 32pp + Sequence Listing; German.

XX CC The present invention provides a number of human immune system associated

CC genes which are modified by the methylation of cytosines. The sequences

CC can be used in the diagnosis and treatment of immune system disorders,

CC including eye diseases such as retinopathy, neovascular glaucoma and

CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid

CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,

CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel

CC diseases. The present sequence is a gene of the invention

XX SQ Sequence 1608 BP; 366 A; 51 C; 421 G; 770 T; 0 U; 0 Other;

Query Match 67.9%; Score 19; DB 6; Length 1608;

Best Local Similarity 81.5%; Pred. No. 1.5e+02;

Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 TCTGGAAGTGCATCCAAATTCAGGTTTC 28

DB 1175 TCTAAGTGCATCCAAATTCAGGTTTC 1149

RESULT 15

AD413316/c

ID AD413316 standard; DNA; 108359 BP.

XX AC AD413316;

XX	FT	06-NOV-2003	(first entry)		FT	variation	/note= "Single nucleotide polymorphism (SNP)"
XX	DT				FT		replace(19074,T)
XX	DE	Human fringe-like secreted protein gene, SEQ ID NO:3.			FT	variation	/*tag= t
XX	XX	Human, secreted protein; glycosyltransferase family; fringe protein;			FT		/note= "Single nucleotide polymorphism (SNP)"
KW	testis; hepatocellular carcinoma; placenta; germinal centre B cell;				FT	variation	replace(20487, .20489,CG)
KW	brain; chromosome 13; drug screening; cancer; cytostatic; gene therapy;				FT		/*tag= u
KW	biochip; transgenic animal; gene; ds.				FT	variation	/note= "Single nucleotide polymorphism (SNP)"
XX	XX	Homo sapiens.			FT		replace(20776,C)
OS	FT				FT	variation	/*tag= v
XX	FT				FT	exon	/note= "Single nucleotide polymorphism (SNP)"
XX	FT				FT		20843, .20919
XX	FT				FT	intron	/*tag= w
XX	FT				FT		20920, .21674
XX	FT				FT	intron	/*tag= x
XX	FT				FT	variation	replace(20923,T)
XX	FT				FT		/*tag= y
XX	FT				FT	variation	/note= "Single nucleotide polymorphism (SNP)"
XX	FT				FT		replace(21188,A)
XX	FT				FT	variation	/*tag= z
XX	FT				FT		/note= "Single nucleotide polymorphism (SNP)"
XX	FT				FT	exon	21675, .21786
XX	FT				FT		/*tag= aa
XX	FT				FT	intron	21787, .34767
XX	FT				FT		/*tag= ab
XX	FT				FT	variation	replace(22922,A)
XX	FT				FT		/*tag= ac
XX	FT				FT	variation	/note= "Single nucleotide polymorphism (SNP)"
XX	FT				FT		replace(23240, .23242,TC)
XX	FT				FT	variation	/*tag= ad
XX	FT				FT		/note= "Single nucleotide polymorphism (SNP)"
XX	FT				FT	variation	replace(23272,G)
XX	FT				FT		/*tag= ae
XX	FT				FT	variation	/note= "Single nucleotide polymorphism (SNP)"
XX	FT				FT		replace(23753,A)
XX	FT				FT	variation	/*tag= af
XX	FT				FT		/note= "Single nucleotide polymorphism (SNP)"
XX	FT				FT	variation	replace(24309,G)
XX	FT				FT		/*tag= ag
XX	FT				FT	variation	/note= "Single nucleotide polymorphism (SNP)"
XX	FT				FT		replace(24333,T)
XX	FT				FT	variation	/*tag= ah
XX	FT				FT		/note= "Single nucleotide polymorphism (SNP)"
XX	FT				FT	variation	replace(24505,A)
XX	FT				FT		/*tag= ai
XX	FT				FT	variation	/note= "Single nucleotide polymorphism (SNP)"
XX	FT				FT		replace(31202,A)
XX	FT				FT	variation	/*tag= aj
XX	FT				FT		/note= "Single nucleotide polymorphism (SNP)"
XX	FT				FT	variation	replace(31739,G)
XX	FT				FT		/*tag= ak
XX	FT				FT	variation	/note= "Single nucleotide polymorphism (SNP)"
XX	FT				FT		replace(31878,G)
XX	FT				FT	variation	/*tag= al
XX	FT				FT		/note= "Single nucleotide polymorphism (SNP)"
XX	FT				FT	variation	replace(32114,C)
XX	FT				FT		/*tag= am
XX	FT				FT	variation	/note= "Single nucleotide polymorphism (SNP)"
XX	FT				FT		replace(33263,G)
XX	FT				FT	variation	/*tag= an
XX	FT				FT		/note= "Single nucleotide polymorphism (SNP)"
XX	FT				FT	variation	replace(34267,G)
XX	FT				FT		/*tag= ao
XX	FT				FT	variation	/note= "Single nucleotide polymorphism (SNP)"
XX	FT				FT	exon	34768, .34904
XX	FT				FT		/*tag= ap
XX	FT				FT	intron	34905, .43035
XX	FT				FT		/*tag= aq
XX	FT				FT	variation	replace(34953, .34955,AT)
XX	FT				FT		/*tag= ar
XX	FT				FT	variation	/note= "Single nucleotide polymorphism (SNP)"
XX	FT				FT		replace(35060,A)
XX	FT				FT	variation	/*tag= as
XX	FT				FT		/note= "Single nucleotide polymorphism (SNP)"

```

FT variation      replace(35112,G)
FT FT            /tag= at
FT FT            /note= "Single nucleotide polymorphism (SNP)"
FT FT            replace(35918,C)
FT FT            /tag= au
FT FT            /note= "Single nucleotide polymorphism (SNP)"
FT FT            replace(35987,G)
FT FT            /tag= av
FT FT            /note= "Single nucleotide polymorphism (SNP)"
FT FT            replace(36052,T)
FT FT            /tag= aw
FT FT            /note= "Single nucleotide polymorphism (SNP)"
FT FT            replace(36466,T)
FT FT            /tag= ax
FT FT            /note= "Single nucleotide polymorphism (SNP)"
FT FT            replace(36567,A)
FT FT            /tag= ay
FT FT            /note= "Single nucleotide polymorphism (SNP)"
FT FT            replace(37241,A)
FT FT            /tag= az
FT FT            /note= "Single nucleotide polymorphism (SNP)"
FT FT            replace(37852,A)
FT FT            /tag= ba
FT FT            /note= "Single nucleotide polymorphism (SNP)"
FT FT            replace(38959,A)
FT FT            /tag= bb
FT FT            /note= "Single nucleotide polymorphism (SNP)"
FT FT            replace(39945,A)
FT FT            /tag= bc
FT FT            /note= "Single nucleotide polymorphism (SNP)"
FT FT            replace(41831,T)
FT FT            /tag= bd
FT FT            /note= "Single nucleotide polymorphism (SNP)"
FT FT            43036. 43099
FT FT            /tag= be
FT FT            43100. 48332
FT FT            /tag= bf
FT FT            replace(43425,A)
FT FT            /tag= bg
FT FT            /note= "Single nucleotide polymorphism (SNP)"
FT FT            replace(43806,T)
FT FT            /tag= bh
FT FT            /note= "Single nucleotide polymorphism (SNP)"
FT FT            48333. 48452
FT FT            /tag= bi
FT FT            48453. 50526
FT FT            /tag= bj
FT FT            replace(49219,C)
FT FT            /tag= bk
FT FT            /note= "Single nucleotide polymorphism (SNP)"
FT FT            replace(49259,A)

```

```

Query Match      67.9%; Score 19; DB 8; Length 108359;
Best Local Similarity 81.5%; Pred. No. 3e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 2 TCTGGAACTGCATCCAAATTCAGGTTTC 28
Db 49691 TCTGGAACTGCAAACAATTACAGATTTC 49665

```

```

Search completed: September 1, 2004, 20:55:46
Job time : 236 secs

```


OM nucleic - nucleic search, using sw model

Run on: September 1, 2004, 17:48:16 ; Search time 47 Seconds
(without alignments)
330.609 Million cell updates/sec

Title: US-10-028-415-2
Perfect score: 28
Sequence: 1 gctgggaactgcattcaattcaggttc 28

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	100.0	28	2	US-08-713-557B-2
2	28	100.0	28	2	US-08-713-557B-14
3	28	100.0	50	2	US-08-713-557B-35
4	28	100.0	2165	4	US-09-685-615B-94
5	17.4	62.1	572	1	US-08-253-155A-15
6	17.4	62.1	1250	4	US-09-023-655-580
7	17.4	62.1	4403785	3	US-09-103-840A-2
8	17.4	62.1	4411529	3	US-09-103-840A-1
9	17.2	61.4	291	4	US-09-540-236-395
10	17.2	61.4	1647	4	US-09-540-236-415
11	17.2	61.4	92407	4	US-09-596-002-36
12	17	60.7	305	4	US-09-313-294A-4746
13	17	60.7	483	4	US-09-328-352-2611
14	17	60.7	722	3	US-08-480-640A-222
15	17	60.7	722	3	US-08-686-968C-222
16	17	60.7	722	3	US-08-488-237A-222
17	17	60.7	722	4	US-08-472-679H-222
18	17	60.7	794	4	US-09-173-300-10
19	17	60.7	3164	3	US-08-686-968C-1
20	17	60.7	12619	4	US-09-616-289-49
21	17	60.7	13187	4	US-09-422-936-61
22	17	60.7	50000	4	US-09-146-053-4
23	17	60.7	65042	4	US-09-784-316-3
24	16.8	60.0	2496	1	US-08-073-384C-2
25	16.8	60.0	2496	1	US-08-254-359A-2
26	16.8	60.0	2496	1	US-08-483-043-2
27	16.8	60.0	2496	1	US-08-481-238-2

28	16.8	60.0	2496	2	US-08-471-066B-2	Sequence 2, Appli
29	16.8	60.0	2496	2	US-08-484-956-2	Sequence 2, Appli
30	16.8	60.0	2496	2	US-08-757-653-2	Sequence 2, Appli
31	16.8	60.0	2496	2	US-08-599-491-2	Sequence 2, Appli
32	16.8	60.0	2496	2	US-08-756-386-2	Sequence 2, Appli
33	16.8	60.0	2496	2	US-08-823-516-2	Sequence 2, Appli
34	16.8	60.0	2496	3	US-08-682-853A-2	Sequence 2, Appli
35	16.8	60.0	2496	3	US-08-759-038-2	Sequence 2, Appli
36	16.8	60.0	2496	3	US-08-758-314-2	Sequence 2, Appli
37	16.8	60.0	2496	4	US-09-350-303-2	Sequence 2, Appli
38	16.8	60.0	2496	4	US-08-520-946-2	Sequence 2, Appli
39	16.8	60.0	2496	4	US-09-684-938-2	Sequence 2, Appli
40	16.8	60.0	2496	4	US-09-308-825A-2	Sequence 2, Appli
41	16.8	60.0	2496	4	US-09-758-282B-2	Sequence 2, Appli
42	16.8	60.0	2496	4	US-09-655-378A-2	Sequence 2, Appli
43	16.8	60.0	2850	4	US-09-620-312D-125	Sequence 125, App
44	16.8	60.0	197496	4	US-09-877-177A-10	Sequence 10, Appli
45	16.6	59.3	2493	3	US-08-945-056-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-713-557B-2
; Sequence 2, Application US/08713557B
; Patent No. 5912168
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Rudert, Fritz
; TITLE OF INVENTION: CD95 REGULATORY GENE SEQUENCES
; TITLE OF INVENTION: AND TRANSCRIPTION FACTORS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 30-AUG-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Speckman, Ann W
; REGISTRATION NUMBER: 31,881
; REFERENCE/DOCKET NUMBER: 11000.1004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-713-557B-2

Query Match 100.0%; Score 28; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTCGGAAGTGCATCCAAATTCAGGTTTC 28

```
Db 1 GTCGGAAGTGCATCCAAATTCAGGTTTC 28
|||||
RESULT 2
US-08-713-557B-14/c
; Sequence 14, Application US/08713557B
; Patent No. 5912168
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Rudert, Fritz
; TITLE OF INVENTION: CD95 REGULATORY GENE SEQUENCES
; TITLE OF INVENTION: AND TRANSCRIPTION FACTORS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/713,557B
; FILING DATE: 30-AUG-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Speckman, Ann W
; REGISTRATION NUMBER: 31,881
; REFERENCE/DOCKET NUMBER: 11000.1004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-713-557B-35
Query Match 100.0%; Score 28; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 0.0003;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGGAAGTGCATCCAAATTCAGGTTTC 28
|||||
Db 1 GTCGGAAGTGCATCCAAATTCAGGTTTC 28
|||||
RESULT 4
US-09-665-615B-94
; Sequence 94, Application US/09665615B
; Patent No. 6653133
; GENERAL INFORMATION:
; APPLICANT: Dean, Nicholas M.
; APPLICANT: Marcusson, Eric G.
; APPLICANT: Wyatt, Jacqueline
; TITLE OF INVENTION: Antisense Modulation of Fas Mediated Signaling
; FILE REFERENCE: ISPH-0502
; CURRENT APPLICATION NUMBER: US/09/665,615B
; CURRENT FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/09/290,640
; PRIOR FILING DATE: 1999-04-12
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 94
; LENGTH: 2165
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1782)...(1813)
US-09-665-615B-94
Query Match 100.0%; Score 28; DB 4; Length 2165;
Best Local Similarity 100.0%; Pred. No. 0.00074;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGGAAGTGCATCCAAATTCAGGTTTC 28
|||||
Db 747 GTCGGAAGTGCATCCAAATTCAGGTTTC 774
|||||
```

```

; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 580:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1250 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: EOSIHET02
; CLONE: 288492
; US-09-023-655-580

Query Match          62.1%; Score 17.4; DB 4; Length 1250;
Best Local Similarity 77.8%; Pred. No. 66;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      2 TCTGGAAGTGCATCCAAATTCAAGTTTC 28
Db      454 TCTGGAACAGGCTCCTCAACTGCAGGCTC 480

RESULT 7
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
; US-09-103-840A-2

Query Match          62.1%; Score 17.4; DB 3; Length 4403765;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      2 TCTGGAAGTGCATCCAAATTCAAGTTTC 28
Db      330484 TCTGGAATCTGAATCTCAACATCAGGTTTC 330458

RESULT 8
US-09-103-840A-1/c

```

; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 62.1%; Score 17.4; DB 3; Length 4411529;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 TCTGGAATCATCAAAATTCAGGTTTC 28
DB 330375 TCTGAATCTGAATCTAATCAGGTTTC 330349

RESULT 9
US-09-540-236-395/c
; Sequence 395, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 395
; LENGTH: 291
; TYPE: DNA
; ORGANISM: M.catarrhalis
US-09-540-236-395

Query Match 61.4%; Score 17.2; DB 4; Length 291;
Best Local Similarity 86.4%; Pred. No. 58;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 GAACTGCATCCAAATTCAGGTT 27
DB 266 GAAATGCATCCAGCTTCAGGTT 245

RESULT 10
US-09-540-236-415
; Sequence 415, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 415
; LENGTH: 1647
; TYPE: DNA
; ORGANISM: M.catarrhalis

Query Match 60.7%; Score 17; DB 4; Length 305;

US-09-540-236-415

Query Match 61.4%; Score 17.2; DB 4; Length 1647;
Best Local Similarity 86.4%; Pred. No. 88;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 GAACTGCATCCAAATTCAGGTT 27
DB 1399 GAAATGCATCCAGCTTCAGGTT 1420

RESULT 11

US-09-596-002-36
; Sequence 36, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 36
; LENGTH: 92407
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte template ID No. 6632636 36
; PUBLICATION INFORMATION:
US-09-596-002-36

Query Match 61.4%; Score 17.2; DB 4; Length 92407;
Best Local Similarity 86.4%; Pred. No. 2.3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 GAACTGCATCCAAATTCAGGTT 27

DB 68664 GAAATGCATCCAGCTTCAGGTT 68685

RESULT 12

US-09-313-294A-4746/c
; Sequence 4746, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Ialugudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 4746
; LENGTH: 305
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700348970H1
; NAME/KEY: unsure
; LOCATION: 47
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-4746

Query Match

```
;
; HYPOTHETICAL: N
; ANTI-SENSE: N
US-08-480-640A-222

Query Match      60.7%; Score 17; DB 3; Length 722;
Best Local Similarity 80.0%; Pred. No. 90;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      2 TCTGGAAGTGCATCCAAATTCAGGT 26
      ||||| ||||| ||||| ||||| |||||
Db      625 TATGGAACAGCATCCCAATTCAGAT 649

RESULT 15
US-08-686-968C-222
; Sequence 222, Application US/08686968C
; Patent No. 6221361
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D.
; TITLE OF INVENTION: Recombinant Swinepox Virus
; FILE REFERENCE: 39119-H/JML
; CURRENT APPLICATION NUMBER: US/08/686,968C
; CURRENT FILING DATE: 1996-07-25
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 222
; LENGTH: 722
; TYPE: DNA
; ORGANISM: Swinepox virus
US-08-686-968C-222

Query Match      60.7%; Score 17; DB 3; Length 722;
Best Local Similarity 80.0%; Pred. No. 90;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      2 TCTGGAAGTGCATCCAAATTCAGGT 26
      ||||| ||||| ||||| ||||| |||||
Db      625 TATGGAACAGCATCCCAATTCAGAT 649

Search completed: September 1, 2004, 21:27:18
Job time : 54 secs
```

```
;
; HYPOTHETICAL: N
; ANTI-SENSE: N
US-08-480-640A-222

Query Match      60.7%; Score 17; DB 4; Length 483;
Best Local Similarity 80.0%; Pred. No. 81;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      2 TCTGGAAGTGCATCCAAATTCAGGT 26
      ||||| ||||| ||||| ||||| |||||
Db      89 TCTGGAAGTGCATCCCAATTCAGGT 65

RESULT 14
US-08-480-640A-222
; Sequence 222, Application US/08480640A
; Patent No. 6033904
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D.
; APPLICANT: Junker, David E.
; TITLE OF INVENTION: Recombinant Swinepox Virus
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,640A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 222:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 722 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 1, 2004, 21:26:22 ; Search time 247 Seconds
(without alignments)
557.894 Million cell updates/sec

Title: US-10-028-415-2
Perfect score: 28
Sequence: 1 gctcggaactgcatacaattcagggttc 28
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 3237270 seqs, 2460713050 residues
Total number of hits satisfying chosen parameters: 6474540

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	28	100.0	28	10	US-09-997-905A-29
C 2	28	100.0	28	14	US-10-028-415-2
C 3	28	100.0	28	14	US-10-028-415-11
C 4	28	100.0	2165	9	US-09-802-669-94
C 5	28	100.0	2165	13	US-10-619-220-94
C 6	28	100.0	2380	9	US-09-834-291-3
C 7	28	100.0	2827	9	US-09-834-291-4
C 8	28	100.0	3212	9	US-09-834-291-1
C 9	28	100.0	28118	13	US-10-282-174-402
C 10	28	100.0	28118	13	US-10-282-174-403
C 11	28	100.0	45121	12	US-09-997-722-10
C 12	19.2	68.6	275449	13	US-10-087-192-520
C 13	19	67.9	1608	15	US-10-311-455-2392
C 14	19	67.9	2171	13	US-10-424-599-11731

C 15	19	67.9	34562	16	US-10-417-476-28	Sequence 28, Appl
C 16	19	67.9	108359	15	US-10-191-807-3	Sequence 3, Appli
C 17	19	67.9	160771	17	US-10-450-826-86	Sequence 86, Appl
C 18	18.8	67.1	289	13	US-10-424-599-115558	Sequence 115558, A
C 19	18.6	66.4	508	10	US-09-918-995-24092	Sequence 24092, A
C 20	18.6	66.4	1597	16	US-10-108-260A-1917	Sequence 1917, Ap
C 21	18.6	66.4	3363	13	US-10-221-278-85	Sequence 85, Appl
C 22	18.6	66.4	3363	16	US-10-291-172-85	Sequence 39, Appl
C 23	18.6	66.4	3414	15	US-10-119-926-39	Sequence 20446, A
C 24	18.4	65.7	204	16	US-10-027-632-20446	Sequence 22293, A
C 25	18.4	65.7	501	10	US-09-918-995-22299	Sequence 347, App
C 26	18.4	65.7	509	9	US-09-978-295A-347	Sequence 347, App
C 27	18.4	65.7	509	9	US-09-978-697-347	Sequence 347, App
C 28	18.4	65.7	509	9	US-09-978-192A-347	Sequence 347, App
C 29	18.4	65.7	509	9	US-09-999-832A-347	Sequence 347, App
C 30	18.4	65.7	509	10	US-09-978-189-347	Sequence 347, App
C 31	18.4	65.7	509	10	US-09-978-608A-347	Sequence 347, App
C 32	18.4	65.7	509	10	US-09-978-585A-347	Sequence 347, App
C 33	18.4	65.7	509	10	US-09-978-191A-347	Sequence 347, App
C 34	18.4	65.7	509	10	US-09-978-403A-347	Sequence 347, App
C 35	18.4	65.7	509	10	US-09-978-564A-347	Sequence 347, App
C 36	18.4	65.7	509	10	US-09-999-833A-347	Sequence 347, App
C 37	18.4	65.7	509	10	US-09-981-915A-347	Sequence 347, App
C 38	18.4	65.7	509	10	US-09-978-824-347	Sequence 347, App
C 39	18.4	65.7	509	10	US-09-918-585A-347	Sequence 347, App
C 40	18.4	65.7	509	10	US-09-978-423A-347	Sequence 347, App
C 41	18.4	65.7	509	10	US-09-978-193A-347	Sequence 347, App
C 42	18.4	65.7	509	10	US-09-999-830A-347	Sequence 347, App
C 43	18.4	65.7	509	10	US-09-978-757A-347	Sequence 347, App
C 44	18.4	65.7	509	10	US-09-978-187B-347	Sequence 347, App
C 45	18.4	65.7	509	10	US-09-978-643A-347	Sequence 347, App

ALIGNMENTS

RESULT 1
US-09-997-905A-29/c
; Sequence 29, Application US/09997905A
; Publication No. US20030074684A1
; GENERAL INFORMATION:
; APPLICANT: Benitec Australia Ltd
; APPLICANT: State of Queensland through its Department of Primary Industries
; TITLE OF INVENTION: Control of Gene Expression
; FILE REFERENCE: W80219470
; CURRENT APPLICATION NUMBER: US/09/997,905A
; CURRENT FILING DATE: 2002-11-30
; PRIOR APPLICATION NUMBER: US 09/100,812
; PRIOR FILING DATE: 1998-06-19
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 28
; TYPE: DNA
; ORGANISM: double-stranded
US-09-997-905A-29

Query Match 100.0%; Score 28; DB 10; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGGAAGTGCATCCAAATTCAGGTTTC 28
|||||
Db 28 GTCGGAAGTGCATCCAAATTCAGGTTTC 1

RESULT 2
US-10-028-415-2
; Sequence 2, Application US/10028415
; Publication No. US20020151063A1
; GENERAL INFORMATION:
; APPLICANT: Lasham, Annette
; APPLICANT: Watson, James D.

; TITLE OF INVENTION: Methods for Modulating Apoptotic Cell

; FILE REFERENCE: 11000.1004c3

; CURRENT APPLICATION NUMBER: US/10/028.415

; CURRENT FILING DATE: 2001-12-20

; PRIOR APPLICATION NUMBER: PCT/NZ01/00286

; PRIOR FILING DATE: 2001-11-28

; PRIOR APPLICATION NUMBER: US 09/724,809

; PRIOR FILING DATE: 2000-11-28

; PRIOR APPLICATION NUMBER: US 09/036,004

; PRIOR FILING DATE: 1998-03-04

; PRIOR APPLICATION NUMBER: US 08/713,557

; PRIOR FILING DATE: 1996-08-30

; SOFTWARE: FastSeq for Windows Version 4.0

; NUMBER OF SEQ ID NOS: 40

; SEQ ID NO 2

; LENGTH: 28

; TYPE: DNA

; ORGANISM: Human

US-10-028-415-2

Query Match 100.0%; Score 28; DB 14; Length 28;

Best Local Similarity 100.0%; Pred. No. 0.0032;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGTGGAACCTGCATCCAAATTCAGGTTTC 28

DB 1 GTCGTGGAACCTGCATCCAAATTCAGGTTTC 28

RESULT 3

US-10-028-415-11/c

; Sequence 11, Application US/10028415

; Publication No. US20020151063A1

; GENERAL INFORMATION:

; APPLICANT: Lasham, Annette

; TITLE OF INVENTION: Methods for Modulating Apoptotic Cell

; FILE REFERENCE: 11000.1004c3

; CURRENT APPLICATION NUMBER: US/10/028.415

; CURRENT FILING DATE: 2001-12-20

; PRIOR APPLICATION NUMBER: PCT/NZ01/00286

; PRIOR FILING DATE: 2001-11-28

; PRIOR APPLICATION NUMBER: US 09/724,809

; PRIOR FILING DATE: 2000-11-28

; PRIOR APPLICATION NUMBER: US 09/036,004

; PRIOR FILING DATE: 1998-03-04

; PRIOR APPLICATION NUMBER: US 08/713,557

; PRIOR FILING DATE: 1996-08-30

; SOFTWARE: FastSeq for Windows Version 4.0

; NUMBER OF SEQ ID NOS: 40

; SEQ ID NO 11

; LENGTH: 28

; TYPE: DNA

; ORGANISM: Human

US-10-028-415-11

Query Match 100.0%; Score 28; DB 14; Length 28;

Best Local Similarity 100.0%; Pred. No. 0.0032;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGTGGAACCTGCATCCAAATTCAGGTTTC 28

DB 28 GTCGTGGAACCTGCATCCAAATTCAGGTTTC 1

RESULT 4

US-09-802-669-94

; Sequence 94, Application US/09802669

; Patent No. US20020004490A1

; GENERAL INFORMATION:

; APPLICANT: Dean, Nicholas M.

; APPLICANT: Marcussen, Eric G.
; APPLICANT: Wyatt, Jacqueline
; APPLICANT: Zhang, Hong
; TITLE OF INVENTION: Antisense Compound Modulation of Fas Mediated Signaling
; FILE REFERENCE: ISPH-545
; CURRENT APPLICATION NUMBER: US/09/802,669
; CURRENT FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: US 09/665,615
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US 09/290,640
; PRIOR FILING DATE: 1999-04-12
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 94
; LENGTH: 2165
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1782)...(1813)
US-09-802-669-94

Query Match 100.0%; Score 28; DB 9; Length 2165;

Best Local Similarity 100.0%; Pred. No. 0.0075;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGTGGAACCTGCATCCAAATTCAGGTTTC 28

DB 747 GTCGTGGAACCTGCATCCAAATTCAGGTTTC 774

RESULT 5

US-10-619-220-94

; Sequence 94, Application US/10619220

; Publication No. US20040033979A1

; GENERAL INFORMATION:

; APPLICANT: Dean, Nicholas M.

; APPLICANT: Marcussen, Eric G.

; APPLICANT: Wyatt, Jacqueline

; APPLICANT: Zhang, Hong

; TITLE OF INVENTION: Antisense Compound Modulation of Fas Mediated Signaling

; FILE REFERENCE: ISPH-545

; CURRENT APPLICATION NUMBER: US/10/619,220

; CURRENT FILING DATE: 2003-07-14

; PRIOR APPLICATION NUMBER: 09/802,669

; PRIOR FILING DATE: 2001-03-01

; PRIOR APPLICATION NUMBER: US 09/665,615

; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: US 09/290,640

; PRIOR FILING DATE: 1999-04-12

; NUMBER OF SEQ ID NOS: 180

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 94

; LENGTH: 2165

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1782)...(1813)

US-10-619-220-94

Query Match 100.0%; Score 28; DB 13; Length 2165;

Best Local Similarity 100.0%; Pred. No. 0.0075;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGTGGAACCTGCATCCAAATTCAGGTTTC 28

DB 747 GTCGTGGAACCTGCATCCAAATTCAGGTTTC 774

RESULT 6

US-09-834-291-3

; Sequence 3, Application US/09834291


```
; Patent No. US20020042064A1
; GENERAL INFORMATION:
; APPLICANT: Krammer, Peter
; APPLICANT: Muller-Schilling, Martina
; APPLICANT: Oren, Moshe
; TITLE OF INVENTION: p53 Binding Areas
; FILE REFERENCE: 4121-122
; CURRENT APPLICATION NUMBER: US/09/834,291
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/DE99/03343
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: DE 198 47 779.1
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2380
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-09-834-291-3

Query Match      100.0%; Score 28; DB 9; Length 2380;
Best Local Similarity 100.0%; Pred. No. 0.0076;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGTGGAAGTGCATCCAAATTCAGGTTTC 28
Db 746 GTCGTGGAAGTGCATCCAAATTCAGGTTTC 773

RESULT 7
US-09-834-291-4
; Sequence 4, Application US/09834291
; Patent No. US20020042064A1
; GENERAL INFORMATION:
; APPLICANT: Krammer, Peter
; APPLICANT: Muller-Schilling, Martina
; APPLICANT: Oren, Moshe
; TITLE OF INVENTION: p53 Binding Areas
; FILE REFERENCE: 4121-122
; CURRENT APPLICATION NUMBER: US/09/834,291
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/DE99/03343
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: DE 198 47 779.1
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2827
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-09-834-291-4

Query Match      100.0%; Score 28; DB 9; Length 2827;
Best Local Similarity 100.0%; Pred. No. 0.0079;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGTGGAAGTGCATCCAAATTCAGGTTTC 28
Db 1193 GTCGTGGAAGTGCATCCAAATTCAGGTTTC 1220

RESULT 8
US-09-834-291-1
; Sequence 1, Application US/09834291
; Patent No. US20020042064A1
; GENERAL INFORMATION:
; APPLICANT: Krammer, Peter
; APPLICANT: Muller-Schilling, Martina
; APPLICANT: Oren, Moshe
; TITLE OF INVENTION: p53 Binding Areas
; FILE REFERENCE: 4121-122
```

```
; CURRENT APPLICATION NUMBER: US/09/834,291
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/DE99/03343
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: DE 198 47 779.1
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3212
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-09-834-291-1

Query Match      100.0%; Score 28; DB 9; Length 3212;
Best Local Similarity 100.0%; Pred. No. 0.0081;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGTGGAAGTGCATCCAAATTCAGGTTTC 28
Db 1192 GTCGTGGAAGTGCATCCAAATTCAGGTTTC 1219

RESULT 9
US-10-282-174-402
; Sequence 402, Application US/10282174
; Publication No. US20030224380A1
; GENERAL INFORMATION:
; APPLICANT: Becker, Kenneth David
; APPLICANT: Velicelabi, Gonul
; APPLICANT: Elliot, Kathryn J.
; APPLICANT: Wang, Xin
; APPLICANT: Tanzi, Rudolph E.
; APPLICANT: Bertram, Lars
; APPLICANT: Saunders, Aleister J.
; APPLICANT: Mullin, Kristina M.
; APPLICANT: Sampson, Andrew Johnson
; APPLICANT: Blacker, Deborah Lynne
; TITLE OF INVENTION: GENES AND POLYMORPHISMS ON CHROMOSOME 10
; TITLE OF INVENTION: ASSOCIATED WITH ALZHEIMER'S DISEASE AND OTHER
; TITLE OF INVENTION: NEURODEGENERATIVE DISEASES
; FILE REFERENCE: 37481-3308
; CURRENT APPLICATION NUMBER: US/10/282,174
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: US 60/339,525
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 60/338,010
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/336,929
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/338,363
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/337,052
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 60/368,919
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 564
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 402
; LENGTH: 28118
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-282-174-402

Query Match      100.0%; Score 28; DB 13; Length 28118;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGTGGAAGTGCATCCAAATTCAGGTTTC 28
Db 1186 GTCGTGGAAGTGCATCCAAATTCAGGTTTC 1213
```

RESULT 10
 US-10-282-174-403
 ; Sequence 403, Application US/10282174
 ; Publication No. US20030224380A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Becker, Kenneth David
 ; APPLICANT: Velicelebi, Gonul
 ; APPLICANT: Elliot, Kathryn J.
 ; APPLICANT: Wang, Xin
 ; APPLICANT: Tanzi, Rudolph E.
 ; APPLICANT: Bertram, Lars
 ; APPLICANT: Saunders, Aleister J.
 ; APPLICANT: Mullin, Kristina M.
 ; APPLICANT: Sampson, Andrew Johnson
 ; APPLICANT: Blacker, Deborah Lynne
 ; TITLE OF INVENTION: GENES AND POLYMORPHISMS ON CHROMOSOME 10
 ; TITLE OF INVENTION: ASSOCIATED WITH ALZHEIMER'S DISEASE AND OTHER
 ; TITLE OF INVENTION: NEURODEGENERATIVE DISEASES
 ; FILE REFERENCE: 37481-3308
 ; CURRENT APPLICATION NUMBER: US/10/282,174
 ; PRIOR FILING DATE: 2002-10-25
 ; PRIOR APPLICATION NUMBER: US 60/339,525
 ; PRIOR FILING DATE: 2001-10-25
 ; PRIOR APPLICATION NUMBER: US 60/338,010
 ; PRIOR FILING DATE: 2001-11-08
 ; PRIOR APPLICATION NUMBER: US 60/336,929
 ; PRIOR FILING DATE: 2001-11-08
 ; PRIOR APPLICATION NUMBER: US 60/338,363
 ; PRIOR FILING DATE: 2001-11-09
 ; PRIOR APPLICATION NUMBER: US 60/337,052
 ; PRIOR FILING DATE: 2001-12-04
 ; PRIOR APPLICATION NUMBER: US 60/368,919
 ; PRIOR FILING DATE: 2002-03-28
 ; NUMBER OF SEQ ID NOS: 564
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 403
 ; LENGTH: 28118
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: allele
 ; LOCATION: 199,213,843,1530,1550,14525,14714,18982,19069,20412,20552,
 ; LOCATION: 23199,23416,24890,26355,29667,3103,5335,5345,6074,9374,9907
 ; OTHER INFORMATION: N is any
 ; FEATURE:
 ; NAME/KEY: allele
 ; LOCATION: 9936,10937,11200,11279,11359,11503,11511,11587,11694,11905,
 ; LOCATION: 12193,12208,12238,18511,18567,18982,19069,20640,21585,22439
 ; OTHER INFORMATION: N is any
 ; FEATURE:
 ; NAME/KEY: allele
 ; LOCATION: 23416,25081,26355,26878,27670
 ; OTHER INFORMATION: N is any
 ; US-10-282-174-403

Query Match 100.0%; Score 28; DB 13; Length 28118;
 Best Local Similarity 100.0%; Pred. No. 0.012;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCTGGAAGTGCATCCAAATTCAGGTT 28
 DB 1186 GTCTGGAAGTGCATCCAAATTCAGGTT 1213

RESULT 11
 US-09-997-722-10
 ; Sequence 10, Application US/09997722
 ; Publication No. US20040072154A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David
 ; APPLICANT: Engelhard, Eric
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
 ; FILE REFERENCE: A-71171/RMS/DCP

; CURRENT APPLICATION NUMBER: US/09/997,722
 ; CURRENT FILING DATE: 2001-11-30
 ; PRIOR APPLICATION NUMBER: US 09/747,377
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: US 09/798,586
 ; PRIOR FILING DATE: 2001-03-02
 ; NUMBER OF SEQ ID NOS: 301
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 10
 ; LENGTH: 45121
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-997-722-10

Query Match 100.0%; Score 28; DB 12; Length 45121;
 Best Local Similarity 100.0%; Pred. No. 0.014;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCTGGAAGTGCATCCAAATTCAGGTT 28
 DB 9186 GTCTGGAAGTGCATCCAAATTCAGGTT 9213

RESULT 12
 US-10-087-192-520/C
 ; Sequence 520, Application US/10087192
 ; Publication No. US20020182586A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David W.
 ; APPLICANT: Engelhard, Eric K.
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
 ; TITLE OF INVENTION: CANCER
 ; FILE REFERENCE: 52945200122
 ; CURRENT APPLICATION NUMBER: US/10/087,192
 ; CURRENT FILING DATE: 2002-03-01
 ; PRIOR APPLICATION NUMBER: US 09/747,377
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: US 09/798,586
 ; PRIOR FILING DATE: 2001-03-02
 ; NUMBER OF SEQ ID NOS: 2059
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 520
 ; LENGTH: 275449
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)-(275449)
 ; OTHER INFORMATION: n = A,T,C or G
 ; US-10-087-192-520

Query Match 68.6%; Score 19.2; DB 13; Length 275449;
 Best Local Similarity 87.5%; Pred. No. 2.7e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 TGGAACTGCATCCAAATTCAGGTT 27
 DB 113532 TGGAACTTATCCAAATTCAGGTT 113509

RESULT 13
 US-10-311-455-2392/C
 ; Sequence 2392, Application US/10311455
 ; Publication No. US20030143606A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENBROCK, Christian
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
 ; TITLE OF INVENTION: cytosine methylation
 ; FILE REFERENCE: 5013.1014
 ; CURRENT APPLICATION NUMBER: US/10/311,455
 ; CURRENT FILING DATE: 2002-12-16

```

; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 2392
; LENGTH: 1608
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2392

Query Match          67.9%; Score 19; DB 15; Length 1608;
Best Local Similarity 81.5%; Pred. No. 1.2e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      2      TCTGGAACTGCATCCAAATTCAGGTTTC 28
Db      1175    TCTAAACTACATCCAAATTCAAATTC 1149

RESULT 14
US-10-424-599-11731/c
; Sequence 11731, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 11731
; LENGTH: 2171
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110600C.1
US-10-424-599-11731

Query Match          67.9%; Score 19; DB 13; Length 2171;
Best Local Similarity 81.5%; Pred. No. 1.3e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      2      TCTGGAACTGCATCCAAATTCAGGTTTC 28
Db      742    TCTTGAAGTGCATCTTATTTCAGGCTC 716

RESULT 15
US-10-417-476-28/c
; Sequence 28, Application US/10417476
; Publication No. US20040002102A1
; GENERAL INFORMATION:
; APPLICANT: Litman, Gary W.
; APPLICANT: Hawke, No. US20040002102A11 A.
; APPLICANT: Yoder, Jeffrey A.
; APPLICANT: Eason, Donna D.
; TITLE OF INVENTION: BIVM (Basic, Immunoglobulin-Like Variable Motif-Containing) Gene,
; TITLE OF INVENTION: Transcriptional Products, and Uses Thereof
; FILE REFERENCE: USF-103X
; CURRENT APPLICATION NUMBER: US/10/417,476
; CURRENT FILING DATE: 2003-04-16
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 28

```

Thu Sep 2 08:32:22 2004

us-10-028-415-2.rnpb

Page 6

Query Match 67.9%; Score 19; DB 16; Length 34562;
Best Local Similarity 81.5%; Pred. No. 2.2e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GTCGGAACTCGATCCAAATTCAGTT 27
|||
Db 19650 GTCGGAATCTACATCCAACTTAAAGTT 19624

Search completed: September 1, 2004, 22:39:27
JOB time : 249 secs

OM nucleic - nucleic search, using sw model

Run on: September 1, 2004, 17:35:32 ; Search time 1821 Seconds
(without alignments)
459.166 Million cell updates/sec

Title: US-10-028-415-2
Perfect score: 28
Sequence: 1 gctggaactgcatccaaattcaggttc 28
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthm:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vri:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20.6	73.6	502	BZ194237	BZ194237 CH230-465
2	20.2	72.1	388	BF941188	BF941188 7d94f07.x
3	20.2	72.1	916	BX372157	BX372157 BX372157
4	20	71.4	464	CB891255	CB891255 EST648224

5	20	71.4	472	14	W24797
6	20	71.4	473	9	AA660188
7	20	71.4	523	10	BF485087
8	20	71.4	525	14	CA718977
9	20	71.4	553	10	BF081044
10	20	71.4	757	29	EX148561
11	20	71.4	845	29	CG972280
12	20	71.4	855	12	BG418680
13	20	71.4	959	13	EX370255
14	19.8	70.7	411	12	BI713285
15	19.8	70.7	530	13	BX526617
16	19.8	70.7	537	28	AZ812626
17	19.8	70.7	689	12	BG866139
18	19.8	70.7	692	12	BI647982
19	19.8	70.7	719	12	BI150689
20	19.8	70.7	779	13	BU611225
21	19.8	70.7	791	12	BI694358
22	19.8	70.7	879	12	BI557142
23	19.8	70.7	892	10	BF181589
24	19.8	70.7	912	13	BQ946032
25	19.8	70.7	927	14	CF583902
26	19.6	70.0	324	13	BQ703424
27	19.6	70.0	398	14	R33603
28	19.6	70.0	460	14	CD482962
29	19.6	70.0	556	29	FR0021171
30	19.6	70.0	611	29	FR0021181
31	19.6	70.0	658	10	BF137351
32	19.2	68.6	275	10	BF560422
33	19.2	68.6	451	28	AQ684637
34	19.2	68.6	527	9	AL796563
35	19.2	68.6	594	14	CB515058
36	19.2	68.6	636	29	CE203423
37	19.2	68.6	655	12	EG591694
38	19.2	68.6	672	14	CF440668
39	19.2	68.6	674	28	BH189737
40	19.2	68.6	674	29	CNS07TGS
41	19.2	68.6	678	28	CC310629
42	19.2	68.6	704	13	CA061797
43	19.2	68.6	725	14	CF441901
44	19.2	68.6	737	14	CF441539
45	19.2	68.6	739	14	CF438864

ALIGNMENTS

RESULT 1
BZ194237
LOCUS BZ194237
DEFINITION CH230-465F5.TJB CHORI-230 Segment 2 Rattus norvegicus genomic clone
ACCESSION BZ194237
VERSION BZ194237.1 GI:23852289
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 502)
AUTHORS Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Riggs,P., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 MboI segment
TITLE Unpublished (1999)
JOURNAL Other GSSs: CH230-465F5.TVB
COMMENT Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

BZ194237 502 bp DNA linear GSS 11-OCT-2002
CH230-465F5.TJB CHORI-230 Segment 2 Rattus norvegicus genomic clone
CH230-465F5, Genomic survey sequence.

Clones are derived from the rat BAC library CHORI-230
(<http://www.chori.org/bacpac/rat230.htm>). For BAC library
availability, please contact Pieter de Jong (pdjong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/orering_information.htm). BAC end
page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 465 row: F column: 5
Seq primer: SP6
Class: BAC ends.

FEATURES

Location/Qualifiers

```
1. 502
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SSHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-465F5"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/notes="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
CHORI-230 Rat (BN/SSHsd/MCW) BAC library produced by
Pieter de Jong"
```

ORIGIN

```
Query Match 73.6%; Score 20.6; DB 28; Length 502;
Best Local Similarity 85.2%; Pred. No. 4.3e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 GTCTGGAACTGCATCCAAATTCAGGTT 27
|||||
DB 379 GTCTGGAAATGCTTCCAAATGCAGTTT 405
|||||
```

RESULT 2

LOCUS

```
DEFINITION BF941188 388 bp mRNA linear EST 22-JAN-2001
IMAGE:3280668 3', mRNA sequence.
```

ACCESSION

```
BF941188
```

VERSION

```
BF941188.1 GI:12358508
```

KEYWORDS

```
EST.
```

SOURCE

```
Organism Homo sapiens (human)
```

REFERENCE

```
1 (bases 1 to 388)
```

AUTHORS

```
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
```

TITLE

```
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
```

JOURNAL

```
Unpublished (1997)
```

COMMENT

```
Contact: Robert Strausberg, Ph.D.
```

```
Email: cgaps-x@mail.nih.gov
```

```
Tissue Procurement: Dr. James Lupski
```

```
CDNA Library Preparation: Lupski Laboratory
```

```
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
```

```
Clone distribution: NCI-CGAP clone distribution information can be
```

```
found through the I.M.A.G.E. Consortium/LLNL at:
```

```
infoimage.llnl.gov
```

```
Seq primer: -40UP from Gibco.
```

FEATURES

Location/Qualifiers

```
1. 388
```

```
/organism="Homo sapiens"
```

```
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"
```

```
/clone="IMAGE:3280668"
```

```
/sex="male"
```

```
/tissue_type="dorsal root ganglia"
```

```
/dev_stage="adult, 36 yr"
```

```
/lab_host="DH10B"
```

```
/clone_lib="Lupski dorsal root ganglion"
```

```
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site 1:
```

```
NotI. Site 2: SalI. cDNA made by oligo-dT priming
```

ORIGIN

```
Query Match 72.1%; Score 20.2; DB 10; Length 388;
Best Local Similarity 88.0%; Pred. No. 5.8e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 TCTGGAATCGCATCCAAATTCAGGT 26
|||||
```

```
DB 228 TCTGGAATCGCATCCAAATTCAGT 252
|||||
```

RESULT 3

LOCUS

```
DEFINITION BX372157 916 bp mRNA linear EST 08-MAY-2003
CDNA clone CS0DK010YH06 3-PRIME, mRNA sequence.
```

ACCESSION

```
BX372157
```

VERSION

```
BX372157.1 GI:30434025
```

KEYWORDS

```
EST.
```

SOURCE

```
Organism Homo sapiens (human)
```

REFERENCE

```
1 (bases 1 to 916)
```

AUTHORS

```
Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
```

TITLE

```
Full-length cDNA libraries and normalization
```

JOURNAL

```
Unpublished (2001)
```

COMMENT

```
Contact: Genoscope
```

```
Genoscope - Centre National de Sequencage
```

```
BP 191 91006 EVRY cedex - France
```

```
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
```

```
Library was constructed by Life Technologies, a division of
```

```
Invitrogen. This sequence belongs to sequence cluster 118.f For
```

```
more information about this cluster, see
```

```
http://www.genoscope.cns.fr/
```

```
http://bin/cluster.cgi?seq=CS0BA1040ZE02\_CS03757\_1acluster=118.f.
```

```
Contact : Feng Liang Email : fliang@lifetech.com URL :
```

```
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
```

```
Faraday Avenue Genoscope sequence ID : CS0BA1040ZE02_CS03757_1.
```

FEATURES

Location/Qualifiers

```
1. 916
```

```
/organism="Homo sapiens"
```

```
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"
```

```
/clone="CS0DK010YH06"
```

```
/cell_type="HELA CELLS COT 25-NORMALIZED"
```

```
/cell_line="HELA"
```

```
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
```

```
/notes="First strand cDNA was primed with a NotI-oligo(dT)
```

```
primer. Five prime end enriched, double-strand cDNA was
```

```
digested with Not I and cloned into the Not I and EcoR V
```

```
sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

```
Query Match 72.1%; Score 20.2; DB 13; Length 916;
Best Local Similarity 88.0%; Pred. No. 7.4e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 4 TGGAACTGCATCCAAATTCAGGTT 28
|||||
```

```
DB 816 TGGAACTGCATCCAAATTCAGGTC 840
|||||
```

RESULT 4

LOCUS

LOCUS
DEFINITION CB891255 464 bp mRNA linear EST 24-APR-2003
EST648224 KV3 Medicago truncatula cDNA clone KV3-49L11, mRNA
sequence.
ACCESSION CB891255
VERSION CB891255.1 GI:30098423
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 464)
AUTHORS VandenBosch,K., Endre,G., Hur,J., Beremand,P., Town,C.D., Van
Aken,S., Uterback,T., Cho,J. and Fraser,C.M.
TITLE ESTs from roots of Medicago truncatula 72 h after Rhizobium
inoculation, 2001
JOURNAL Unpublished (2001)
COMMENT Contact: VandenBosch K
Department of Plant Biology
University of Minnesota
220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
Tel: 612 624 2755
Fax: 612 625 1738
Email: kvandenb@csb.umn.edu
TIGR sequence name: MTECO66TK
More information is available at: www.medicago.org
Seq primer: SKmcd (CTA GAA CTA gtg gat CC).
FEATURES
source Location/Qualifiers
1..464
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="genotype Al7"
/db_xref="taxon:3880"
/clone="KV3-49L11"
/tissue types="Seedling roots"
/dev_stage="3 days post-inoculation with Sinorhizobium
meliloti"
/lab_host="E. coli strain XL0LR"
/clone_lib="KV3"
/note="Vector: pBluescript SK -; Site 1: EcoRI; Site 2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unizap XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-assist
helper phage and propagated in XL0LR cells."
ORIGIN
Query Match 71.4%; Score 20; DB 14; Length 464;
Best Local Similarity 82.1%; Pred. No. 7.4e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 GTCGTGGAACCTGCATCCAAATTCAGGTTTC 28
Db 135 GGCTTGCACTGCTCATCCAAATACAGCTTC 162
RESULT 5
W24797
LOCUS W24797 472 bp mRNA linear EST 11-OCT-1996
DEFINITION zcf5f05.r1 Soares fetal heart NBHHL19W Homo sapiens cDNA clone
IMAGE:327201 5', mRNA sequence.
ACCESSION W24797
VERSION W24797.1 GI:1303640
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 472)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfig,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 2873 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 311.
FEATURES
source Location/Qualifiers
1..472
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1261403"
/db_xref="taxon:9606"
/clone="IMAGE:327201"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal heart NBHHL19W"
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCATCTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NBHHL19W."

Gilbert Biology, Stanford, CA 94305-5020, USA
 Tel: 650 723 3232
 Fax: 650 725 8309
 Email: fa.srl@forsythe.stanford.edu
 Seq primer: 13.

FEATURES

source

Location/Qualifiers
 1..473
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /cultivar="Jemalong"
 /db_xref="taxon:3880"
 /tissue_type="Root hairs & tips"
 /dev_stage="2-3 day old seedlings"
 /clone_lib="McRHP"
 /note="Organ: Root; Vector: pBK-CMV; Site 1: EcoRI;
 Site 2: XhoI; cDNA was synthesized from a pooled mRNA prep
 from elongating root hairs (30% w/w) and 2-3cm root tips
 (70% w/w). XhoI-oligo-dT linker-primer and EcoRI
 adaptors were used. cDNAs were cloned unidirectionally
 into lambda ZAP Express (Stratagene), amplified, and
 mass-excised into pBK-CMV vector plasmids. More
 information is available at <http://bio-SRLs.stanford.edu>."

ORIGIN

Query Match 71.4%; Score 20; DB 9; Length 473;
 Best Local Similarity 82.1%; Pred. No. 7.4e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GTCCTGGAACTGCATCCCAAAATTCAGGTTTC 28

Db 138 GGCTTGACATGCATCCCAAAATACAGCTTC 165

RESULT 7

BF485087

LOCUS

DEFINITION WHE1793_C05_F0925 Wheat pre-anthesis spike cDNA library Triticum
 aestivum cDNA clone WHE1793_C05_F09, mRNA sequence.

VERSION

BF485087

SOURCE

EST.

Triticum aestivum (bread wheat)

ORGANISM

Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.

REFERENCE

1

(bases 1 to 523)

Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D.,
 Han, P.S., Haia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J.,
 Seaton, C.L. and Tong, J.C.The structure and function of the expressed portion of the wheat
 genomes - Pre-anthesis spike cDNA library

Unpublished (2000)

Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific
 West Area, Western Regional Research Center

800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105595773

Fax: 5105595818

Email: olinda@nwp.usda.govSequence have been trimmed to remove vector sequence and low
 quality sequence with phred score less than 20

Seq primer: Stratagene SK primer.

Location/Qualifiers

1..523

/organism="Triticum aestivum"

/mol_type="mRNA"

/cultivar="Chinese Spring"

/db_xref="taxon:4565"

/clone="WHE1793_C05_F09"

/tissue_type="Spike before anthesis"

/dev_stage="Adult plant"

/lab_host="E. coli SOLR"

FEATURES

source

Query Match 71.4%; Score 20; DB 14; Length 525;
 Best Local Similarity 82.1%; Pred. No. 7.7e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GTCCTGGAACTGCATCCCAAAATTCAGGTTTC 28

Db 407 GTCTTGAACAGCATCCCAAGCTTTAGGTTTC 434

RESULT 9

/clone_lib="Wheat pre-anthesis spike cDNA library"
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
 Site 1: EcoRI; Site 2: XhoI; Plants were grown in the
 greenhouse. Whole spike with awns trimmed, white, green
 and yellow anther were collected and total RNA, and
 poly(A) RNA were prepared, a cDNA library was made, and
 the cDNA clones were in vivo excised to give pBluescript
 phagemids in the TJ Close lab (Choi, Close, Fenton) at
 the University of California, Riverside. Plasmid DNA
 preparations and DNA sequencing were performed in the OD
 Anderson lab (all other authors)."

ORIGIN

Query Match 71.4%; Score 20; DB 10; Length 523;
 Best Local Similarity 82.1%; Pred. No. 7.7e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GTCCTGGAACTGCATCCCAAAATTCAGGTTTC 28

Db 83 GACTGGAACTGCTCCCAATATCAGGTTTC 110

RESULT 8

CA718977

LOCUS

DEFINITION

Wkm2n.pk002.06 wkm2n Triticum aestivum cDNA clone wkm2n.pk002.06 5',
 end, mRNA sequence.

ACCESSION

CA718977

VERSION

CA718977.1

KEYWORDS

GI:25440770

SOURCE

Triticum aestivum (bread wheat)

ORGANISM

Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.

REFERENCE

1

(bases 1 to 525)

Tingey, S.V., Powell, W., Wolters, P., Dolan, M., Hainey, C., Yuan, Z.,
 Miao, G., Caraher, N. and Hanafey, M.K.

DuPont Wheat cDNA Sequence

Unpublished (2002)

Contact: Scott V. Tingey

Crop Genetics

E. I. DuPont de Nemours and Company

1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA

Tel: 302-631-2602

Fax: 302-631-2607

Email: Scott.V.Tingey@USA.dupont.com

Seq primer: M13.

Location/Qualifiers

1..525

/organism="Triticum aestivum"

/mol_type="mRNA"

/cultivar="Spring wheat"

/db_xref="taxon:4565"

/clone="wkm2n.pk002.06"

/tissue_type="kernel"

/lab_host="DH108"

/clone_lib="wkm2n"

/note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2:
 XhoI; Wheat (Triticum aestivum L.) kernel malted 175 hours
 at 4 C, normalized"

ORIGIN

Query Match 71.4%; Score 20; DB 14; Length 525;
 Best Local Similarity 82.1%; Pred. No. 7.7e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GTCCTGGAACTGCATCCCAAAATTCAGGTTTC 28

Db 407 GTCTTGAACAGCATCCCAAGCTTTAGGTTTC 434

BF081044/c
LOCUS BF081044 553 bp mRNA linear EST 18-OCT-2000
DEFINITION 23889 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BF081044
VERSION BF081044.1 GI:10874874
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 553)
AUTHORS Fahrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,
Vallet,J., Wise,T., Rohrer,G.A., Pertea,G., Sultana,R.,
Quackenbush,J. and Keefe,J.W.
TITLE Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly
JOURNAL Mamm. Genome 13 (8), 475-478 (2002)
MEDLINE 22213789
PubMed 12226715
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCAGTCAGCAG
Plate: 54 row: P column: 4
Seq primer: ATTAGTGACATCATAG.
Location/Qualifiers
1..553
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 2P1G"
/note="Vector: pCMV SPORT6; Site.1: NotI; Site.2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
ORIGIN
Query Match 71.4%; Score 20; DB 10; Length 553;
Best Local Similarity 82.1%; Pred. No. 7.8e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 GTCCTGGAAGTTCATCAAAATTCAGGTTTC 28
DB 392 GTGTGGAAGTTCATCAAAATTCAGGTTTC 365
RESULT 10
LOCUS BX148561/c 757 bp DNA linear GSS 13-MAR-2003
DEFINITION Danio rerio genomic clone DKEY-101C6, genomic survey sequence.
ACCESSION BX148561
VERSION BX148561.1 GI:27979989
KEYWORDS GSS.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 757)
AUTHORS Humphray,S.J., Huckle,E. and Durham,J.L.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Unpublished

This sequence was generated from the T7 end of BAC 101C6. 101C6 is
part of the Daniokey BAC Library created by R. Piasterk and N.V.
Keygene. Further details:
http://www.sanger.ac.uk/Projects/D_rerio/.

FEATURES
source
1..757
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-101C6"
/tissue_type="Testis"
/note="Vector pIndigoBAC-536"

ORIGIN
Query Match 71.4%; Score 20; DB 29; Length 757;
Best Local Similarity 82.1%; Pred. No. 8.5e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 GTCCTGGAAGTTCATCAAAATTCAGGTTTC 28
DB 347 GTCCGGAAGTTCCTTCTTTCAGGTTTC 320
RESULT 11
CG972280
LOCUS MBEX48TR mth2 Medicago truncatula genomic clone 16H24, genomic
DEFINITION survey sequence.
ACCESSION CG972280
VERSION CG972280.1 GI:39898059
KEYWORDS GSS.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 845)
AUTHORS Town,C.D., Shetty,J., Koo,H. and Feldblyum,T.F.
TITLE Sequencing of BAC ends from Medicago truncatula
JOURNAL Unpublished (2003)
COMMENT Other GSSs: MBEX48TRFB
Contact: Chris Town
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
Seq primer: CAGGAACAGCTATGAC
Class: BAC ends.
Location/Qualifiers
1..845
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="16H24"
/clone_lib="mth2"
/note="Vector: pBelobAC11; Site.1: HindIII; Site.2:
HindIII; Cook, D.R. and Kim, D.J, unpublished"

ORIGIN
Query Match 71.4%; Score 20; DB 29; Length 845;
Best Local Similarity 82.1%; Pred. No. 8.8e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 GTCCTGGAAGTTCATCAAAATTCAGGTTTC 28
DB 313 GGCTTGCAATGATCAAAATTCAGGTTTC 340
RESULT 12
BG418680/c

LOCUS BG418680 855 bp mRNA linear EST 23-OCT-2001
 DEFINITION HVSNK0023008f Hordeum vulgare testa/pericarp EST library
 HVSNK0013 (normal) Hordeum vulgare subsp. vulgare cDNA clone
 HVSNK0023008f, mRNA sequence.
 ACCESSION BG418680
 VERSION BG418680.2 GI:16333779
 SOURCE EST.
 ORGANISM Hordeum vulgare subsp. vulgare
 Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooideae; Triticeae; Hordeum.
 1 (bases 1 to 855)
 Weng, R., Close, T.J., Kleinhofs, A., Wise, R., Kannangara, G., von
 Wetstein, D., Akhunov, E., Chin, A., Choi, D.W., Fenton, R.D.,
 Kianian, P., Otto, C., Simons, K., Zhang, D., Begum, D., Frisch, D.,
 Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Oates, R. and
 Main, D.
 TITLE Development of a genetically and physically anchored EST resource
 for barley genomics: Morex testa/pericarp cDNA library
 JOURNAL Unpublished (2001)
 COMMENT On Mar 13, 2001 this sequence version replaced gi:13324231.
 Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Total hq bases = 507
 Seq primer: AATTACCTCCTCACTAAAGGG
 High quality sequence stop: 568.
 Location/Qualifiers
 1..855
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="Morex"
 /sub_species="vulgare"
 /db_xref="taxon:112509"
 /clone="HVSMEK0023008f"
 /tissue_type="testa/pericarp"
 /lab_host="TJC121"
 /clone_lib="Hordeum vulgare testa/pericarp EST library
 HVSNK0013 (normal)"
 /notes="vector: lambdaZAP; Site1: EcoRI; Site 2: XhoI;
 Plants were raised from seeds in a Controlled Environments
 growth chamber maintained in continuous light at 18°C, and
 testa and pericarp were dissected from developing kernels
 at Washington State University, Pullman, WA (Kannangara,
 von Wetstein). Total RNA was prepared, poly(A) RNA was
 purified, one cDNA library was made, and 1 million pfu
 were in vivo excised to give pBluescript SK(-) cDNA
 phagemids in the TJ Clone lab at the University of
 California, Riverside (Akhunov, Chin, Choi, Close, Fenton,
 Kianian, Otto, Simons, Zhang). Phagemids were plated and
 picked at the Clemson University Genomics Institute (CUGI)
 (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA
 preparations, DNA sequencing and sequence analysis were
 performed at CUGI (Wing, Yu, Frisch, Henry, Simmons,
 Oates, Rambo, Main). The sequence has been trimmed to
 remove vector sequence and contains a minimum of 100 bases
 of phred value 20 or above. For more details on library
 preparation and sequence analysis see
 http://www.genome.clemson.edu/projects/barley. To order
 this clone see http://www.genome.clemson.edu/orders Also
 see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
 Genetically and physically anchored EST resources for
 barley genomics. Barley Genetics Newsletter 31:29-30.
 (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"

Query Match 71.4%; Score 20; DB 12; Length 855;
 Best Local Similarity 82.1%; Pred. No. 8.8e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCTGTGGAAGTGCATCCAAATTCAGGTTTC 28
 |||||
 DB 769 GCTGTGGAACCGCACATAAATTAAGGTTTC 742
 |||||

RESULT 13
 LOCUS BX370255/c
 DEFINITION BX370255 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 clone CS0DI087YE21 5-PRIME, mRNA sequence.
 ACCESSION BX370255
 VERSION BX370255.1 GI:30453891
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 959)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 245.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0BAG040ZF02 CS03758_1&cluster=245.r.
 Contact : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0BAG040ZF02_CS03758_1.
 Location/Qualifiers
 1..959
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DI087YE21"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo (dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match 71.4%; Score 20; DB 13; Length 959;
 Best Local Similarity 82.1%; Pred. No. 9.1e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCTGTGGAAGTGCATCCAAATTCAGGTTTC 28
 |||||
 DB 715 GCCTGGAACTGCTTCCCACTTGAGGTTTC 688
 |||||

RESULT 14
 LOCUS BX1713285
 DEFINITION BX1713285 Mus musculus Normalized Mixed Mouse Pancreas 1 NI-MMS1 Mus
 musculus cDNA clone IMAGE:5660619 5', mRNA sequence.
 ACCESSION BX1713285
 VERSION BX1713285.1 GI:15688980
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 411)
 Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
 Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,

Query Match 71.4%; Score 20; DB 12; Length 855;
 Best Local Similarity 82.1%; Pred. No. 8.8e+02;

This page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 1, 2004, 19:43:07 ; Search time 1400 Seconds
(without alignments)
866.861 Million cell updates/sec

Title: US-10-028-415-2
Perfect score: 28
Sequence: 1 gcttgaactgcattccaaattcaggttc 28

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 5940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	28	100.0	28	6	AR071972	AR071972 Sequence
2	28	100.0	28	6	AR071984	AR071984 Sequence
3	28	100.0	50	6	AR072005	AR072005 Sequence
4	28	100.0	1608	9	HSFASX1	X82279 H.sapiens F
5	28	100.0	1877	9	HSAPT1	X81335 H.sapiens A
6	28	100.0	2165	6	AR432294	AR432294 Sequence
7	28	100.0	2165	9	HUMFAS	D31968 Homo sapien
8	28	100.0	2344	9	HSCD955FR	X87625 H.sapiens C
9	28	100.0	2380	6	BD237717	BD237717 p53 bindi
10	28	100.0	2380	6	AX026091	AX026091 Sequence
11	28	100.0	2827	6	BD237718	BD237718 p53 bindi
12	28	100.0	2827	6	AX026092	AX026092 Sequence
13	28	100.0	3212	6	BD237715	BD237715 p53 bindi
14	28	100.0	3212	6	AX026089	AX026089 Sequence
15	28	100.0	28339	9	AY450925	AY450925 Homo sapi
16	28	100.0	45121	6	AX695635	AX695635 Sequence
17	28	100.0	187313	9	AL157394	AL157394 Human DNA
18	21.6	77.1	197772	2	AC145295	AC145295 Pongo pyg
19	21.6	77.1	205616	2	AC146844	AC146844 Pongo pyg
20	21.2	75.7	4569	5	XLM181	Y08295 X.laevis MR
21	21.2	75.7	121355	9	AL358943	AL358943 Human DNA
22	21.2	75.7	173386	2	AL353776	AL353776 Homo sapi
23	20.8	74.3	75820	2	AC101006	AC101006 Mus muscu
24	20.8	74.3	100000	9	AP000069	AP000069 Homo sapi
25	20.8	74.3	100364	9	AC007590	AC007590 Homo sapi
26	20.8	74.3	160190	2	AC119202	AC119202 Mus muscu
27	20.8	74.3	165610	2	AC134527	AC134527 Mus muscu
28	20.8	74.3	205831	2	AC074192	AC074192 Homo sapi
29	20.8	74.3	208375	2	AC130246	AC130246 Homo sapi
30	20.8	74.3	208375	2	AC130246	AC130246 Homo sapi
31	20.8	74.3	211750	9	AP006306	AP006306 Homo sapi
32	20.8	74.3	230563	2	AC094600	AC094600 Rattus no
33	20.6	73.6	280325	2	AC112293	AC112293 Rattus no
34	20.6	73.6	323304	2	AC108248	AC108248 Rattus no
35	20.2	72.1	158830	2	AC027693	AC027693 Homo sapi
36	20.2	72.1	177187	2	AC096848	AC096848 Papio anu
37	20.2	72.1	187321	2	AC092510	AC092510 Papio anu
38	20.2	72.1	210623	2	AC113027	AC113027 Mus muscu
39	20	71.4	8193	6	AX537570	AX537570 Sequence
40	20	71.4	8193	6	AX537571	AX537571 Sequence
41	20	71.4	16253	9	AY135329	AY135329 Homo sapi
42	20	71.4	126779	9	AL603882	AL603882 Human DNA
43	20	71.4	129779	9	AL359457	AL359457 Human DNA
44	20	71.4	142000	2	EX664725	EX664725 Homo sapi
45	20	71.4	145937	2	AL355600	AL355600 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS AR071972 28 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 2 from patent US 5912168.
ACCESSION AR071972
VERSION AR071972.1 GI:7222860
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 28)
AUTHORS Watson,J.D. and Rudert,F.
TITLE CD95 regulatory gene sequences
JOURNAL Patent: US 5912168-A 2 15-JUN-1999;
FEATURES Location/Qualifiers

Pred. No. is the number of results predicted by chance to have a


```
misc_feature
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="10"
  /map="10q23"
  /clone="CAPO-1"
  /clone_lib="cosmid, pcos2EMBL"
  151..261
  /notes="pyrimidine rich region"
gene
  join(1653..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..613,X81340.1:79..161,
X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
mRNA
  join(1653..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..613,X81339.1:766..828,
X81340.1:79..161,X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HsmRNA1"
  1653..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HsmRNA1
  /label=exib
mRNA
  join(1670..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..613,X81339.1:766..828,
X81340.1:79..161,X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HsmRNA2"
  1670..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HsmRNA2
  /label=exib
mRNA
  join(1689..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..613,X81339.1:766..828,
X81340.1:79..161,X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HsmRNA3"
  1689..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HsmRNA3
  /label=exic
mRNA
  join(1704..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..613,X81339.1:766..828,
X81340.1:79..161,X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HsmRNA4"
  1704..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HsmRNA4
  /label=exid
mRNA
  join(1707..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..613,X81339.1:766..828,
X81340.1:79..161,X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HsmRNA5"
  1707..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HsmRNA5
  /label=exie
mRNA
  join(1708..1810,X81336.1:155..320,X81337.1:146..283,
X81338.1:145..253,X81339.1:552..613,X81339.1:766..828,
X81340.1:79..161,X81341.1:255..279,X81342.1:103..1757)
  /gene="APT"
  /product="APO-1"
  /label="HsmRNA6"
  1708..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HsmRNA6
  /label=exif

exon
  1708..1810
  /gene="APT"
  /number=1
  /usedin=X81335:HsmRNA6
  /label=exif

ORIGIN
  Query Match 100.0%; Score 28; DB 9; Length 1877;
  Best Local Similarity 100.0%; Pred. No. 0.0081;
  Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

  QY 1 GCTCGAACTGCATCCAAATTCAGGTTTC 28
  Db 746 GTCGAACTGCATCCAAATTCAGGTTTC 773

RESULT 6
LOCUS AR432294 2165 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 94 from patent US 6653133.
ACCESSION AR432294
VERSION AR432294.1 GI:40194567
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
  1 (bases 1 to 2165)
  AUTHORS Dean,N.M., Marcussen,E.G. and Wyatt,J.
  TITLE Antisense modulation of Fas mediated signaling
  JOURNAL Patent: US 6653133-A 94 25-NOV-2003;
  FEATURES
    Location/Qualifiers
      source
        1..2165
        /organism="unknown"
        /mol_type="genomic DNA"

ORIGIN
  Query Match 100.0%; Score 28; DB 6; Length 2165;
  Best Local Similarity 100.0%; Pred. No. 0.0083;
  Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

  QY 1 GCTCGAACTGCATCCAAATTCAGGTTTC 28
  Db 747 GTCGAACTGCATCCAAATTCAGGTTTC 774

RESULT 7
LOCUS HUMFAS 2165 bp DNA linear PRI 07-FEB-2003
DEFINITION Homo sapiens gene for Fas antigen, partial cds and promoter region.
ACCESSION D31968
VERSION D31968.1 GI:961455
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  AUTHORS Wada,N., Matsumura,M., Ohba,Y., Kobayashi,N., Takizawa,T. and
  Nakanishi,Y.
  TITLE Transcription stimulation of the Fas-encoding gene by nuclear
  factor for interleukin-6 expression upon influenza virus infection
  JOURNAL J. Biol. Chem. 270 (30), 18007-18012 (1995)
  MEDLINE 95355401
  PUBMED 7543095
  REFERENCE
    2 (bases 1 to 2165)
  AUTHORS Nakanishi,Y.
  TITLE Direct Submission
  JOURNAL Submitted (28-JUN-1994) Yoshinobu Nakanishi, Kanazawa University,
  Faculty of Pharmaceutical Sciences, 13-1 Takara-machi, Kanazawa,
  Ishikawa 920-0934, Japan (E-mail:nakanaka@db.s.p.kanazawa-u.ac.jp,
  Tel:81-76-234-4424, Fax:81-76-234-4480)
  FEATURES
    Location/Qualifiers
      source
        1..2165
```

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="p7"
/tissue_type="blood"
79..87
/bound_moiety="NF-IL6"
276..284
/bound_moiety="NF-IL6"
481..489
/bound_moiety="NF-IL6"
772..780
/bound_moiety="NF-IL6"
934..940
/bound_moiety="AP-1"
948..955
/bound_moiety="Ets"
1298..1306
/bound_moiety="NF-IL6"
1436..1813
/number=1
1436..1781
1782..1813
/codon_start=1
/product="Fas antigen"
/protein_id="BAA20850.1"
/db_xref="GI:4431150"
/translation="MLGIWTLPLV"
1814..22165
/number=1

ORIGIN
Query Match 100.0%; Score 28; DB 9; Length 2165;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCGGAACTGCATCCAAATTCAGGTTTC 28
|||||
Db 747 GTCGGAACTGCATCCAAATTCAGGTTTC 774

RESULT 8
HSCD955FR HSCD955FR 2344 bp DNA linear PRI 05-FEB-1997
DEFINITION H.sapiens CD95 gene 5' flanking region.
ACCESSION X87625
VERSION X87625.1 GI:902311
KEYWORDS beta interferon; CD95 gene; silencer.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Rudert,F., Visser,E., Forbes,L., Lindridge,E., Wang,Y. and
Watson,J.
TITLE Identification of a silencer, enhancer, and basal promoter region
JOURNAL DNA Cell Biol. 14 (11), 931-937 (1995)
MEDLINE 96069539
PUBMED 7576179
REFERENCE 2 (bases 1 to 2344)
AUTHORS Rudert,F.H.
TITLE Direct Submission
JOURNAL Submitted (26-MAY-1995) F.H. Rudert, Genesis Research &
Development, Corporation Ltd., PO Box 50, Auckland, NEW ZEALAND
COMMENT Overlaps with X81335, & X82279-X82286.
Location/Qualifiers
1..2344
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="q24.1"

FEATURES
source
ORIGIN
Query Match 100.0%; Score 28; DB 9; Length 2344;
Best Local Similarity 100.0%; Pred. No. 0.0084;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCGGAACTGCATCCAAATTCAGGTTTC 28
|||||
Db 746 GTCGGAACTGCATCCAAATTCAGGTTTC 773

RESULT 10
AX026091 AX026091 2380 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 3 from Patent DE19847779.
ACCESSION AX026091
VERSION AX026091.1 GI:10187522
KEYWORDS Homo sapiens (human)
SOURCE

```

```

/tissue_type="placenta"
/germline
564..1337
/note="silencer"
717..801
/note="beta interferon silencer B motifs"
complement(1237..1244)
/note="lysozyme silencer 1"
1338..1919
1920..2344
/note="basal promoter"

ORIGIN
Query Match 100.0%; Score 28; DB 9; Length 2344;
Best Local Similarity 100.0%; Pred. No. 0.0084;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCGGAACTGCATCCAAATTCAGGTTTC 28
|||||
Db 1310 GTCGGAACTGCATCCAAATTCAGGTTTC 1337

RESULT 9
BD237717 BD237717 2380 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION p53 binding region.
ACCESSION BD237717
VERSION BD237717.1 GI:33047487
KEYWORDS JP 2002527108-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Krammer,P., Schilling,M.M. and Oren,M.
TITLE p53 binding region
JOURNAL DEUTSCHES KREBSFORSCHUNGSZENTRUM STIFTUNG DES OFFENTLICHEN RECHTS
PATENT: JP 2002527108-A 3 27-AUG-2002;
COMMENT CS Homo sapiens (human)
PN JP 2002527108-A/3
PD 27-AUG-2002
PP 18-OCT-1999 JP 2000577293
PR 16-OCT-1998 DE 198 47 779.1
PI PETER KRAMMER,MARTINA MUELLER SCHILLING,MOSHE OREN PC
C12N15/09,C12Q1/66,C12Q1/68,G01N33/566,C12N15/00 CC p53 binding
region
FH Key Location/Qualifiers
FT source
Location/Qualifiers
1..2380
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source
ORIGIN
Query Match 100.0%; Score 28; DB 6; Length 2380;
Best Local Similarity 100.0%; Pred. No. 0.0084;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCGGAACTGCATCCAAATTCAGGTTTC 28
|||||
Db 746 GTCGGAACTGCATCCAAATTCAGGTTTC 773

RESULT 10
AX026091 AX026091 2380 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 3 from Patent DE19847779.
ACCESSION AX026091
VERSION AX026091.1 GI:10187522
KEYWORDS Homo sapiens (human)
SOURCE

```


ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Mueller-Schilling, M., Krammer, P. and Oren, M.
Novel receptor dna useful for identifying apoptosis-modulating
substances potentially useful for cancer chemotherapy
Patent: DE 19847779-C 3 03-FEB-2000;
JOURNAL DEUTSCHES KREBSFORSCH (DE)
FEATURES
source
1. .2380
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 28; DB 6; Length 2380;
Best Local Similarity 100.0%; Pred. No. 0.0084;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTCGGAACTGCATCCAAATTCAGGTTTC 28
Db 746 GTCGGAACTGCATCCAAATTCAGGTTTC 773
RESULT 11
BD237718 2827 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION p53 binding region.
ACCESSION BD237718
VERSION BD237718.1 GI:33047488
KEYWORDS JP 2002527108-A/4.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2827)
Krammer, P., Schilling, M.M. and Oren, M.
p53 binding region
Patent: JP 2002527108-A 4 27-AUG-2002;
JOURNAL DEUTSCHES KREBSFORSCHUNGSZENTRUM STIFTUNG DES OFFENTLICHEN RECHTS
COMMENT OS Homo sapiens (human)
PN JP 2002527108-A/4
PD 27-AUG-2002
PF 18-OCT-1999 JP 2000577293
PR 16-OCT-1998 DE 198 47 779.1
PI PETER KRAMMER, MARTINA MUELLER SCHILLING, MOSHE OREN PC
C12N15/09, C12Q1/66, C12Q1/68, G01N33/566, C12N15/00 CC p53 binding
region
FH Key Location/Qualifiers
FT source /organism="Homo sapiens (human)".
FEATURES
source
1. .2827
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 28; DB 6; Length 2827;
Best Local Similarity 100.0%; Pred. No. 0.0086;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTCGGAACTGCATCCAAATTCAGGTTTC 28
Db 1193 GTCGGAACTGCATCCAAATTCAGGTTTC 1220
RESULT 12
AX026092 2827 bp DNA linear PAT 16-SEP-2000
LOCUS
DEFINITION Sequence 4 from Patent DE19847779.
ACCESSION AX026092

VERSION AX026092.1 GI:10187523
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Mueller-Schilling, M., Krammer, P. and Oren, M.
Novel receptor dna useful for identifying apoptosis-modulating
substances potentially useful for cancer chemotherapy
Patent: DE 19847779-C 4 03-FEB-2000;
JOURNAL DEUTSCHES KREBSFORSCH (DE)
FEATURES
source
1. .2827
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 28; DB 6; Length 2827;
Best Local Similarity 100.0%; Pred. No. 0.0086;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTCGGAACTGCATCCAAATTCAGGTTTC 28
Db 1193 GTCGGAACTGCATCCAAATTCAGGTTTC 1220
RESULT 13
BD237715 3212 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION p53 binding region.
ACCESSION BD237715
VERSION BD237715.1 GI:33047485
KEYWORDS JP 2002527108-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 3212)
Krammer, P., Schilling, M.M. and Oren, M.
p53 binding region
Patent: JP 2002527108-A 1 27-AUG-2002;
JOURNAL DEUTSCHES KREBSFORSCHUNGSZENTRUM STIFTUNG DES OFFENTLICHEN RECHTS
COMMENT OS Homo sapiens (human)
PN JP 2002527108-A/1
PD 27-AUG-2002
PF 18-OCT-1999 JP 2000577293
PR 16-OCT-1998 DE 198 47 779.1
PI PETER KRAMMER, MARTINA MUELLER SCHILLING, MOSHE OREN PC
C12N15/09, C12Q1/66, C12Q1/68, G01N33/566, C12N15/00 CC p53 binding
region
FH Key Location/Qualifiers
FT source /organism="Homo sapiens (human)".
FEATURES
source
1. .3212
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 28; DB 6; Length 3212;
Best Local Similarity 100.0%; Pred. No. 0.0087;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTCGGAACTGCATCCAAATTCAGGTTTC 28
Db 1192 GTCGGAACTGCATCCAAATTCAGGTTTC 1219
RESULT 14
AX026089

```

LOCUS       AX026089               3212 bp      DNA      linear      PAT 16-SEP-2000
DEFINITION   Sequence 1 from Patent DE19847779.
ACCESSION    AX026089
VERSION      AX026089.1  GI:10187520
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1
AUTHORS      Mueller-Schilling, M., Kramer, P. and Oren, M.
TITLE        Novel receptor dna useful for identifying apoptosis-modulating
             substances potentially useful for cancer chemotherapy
JOURNAL      Patent: DE 19847779-C 1 03-FEB-2000;
DESCRIPTORS  DEUTSCHES KREBSFORSCH (DE)
FEATURES     Location/Qualifiers
             1..3212
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"
ORIGIN
Query Match      100.0%; Score 28; DB 6; Length 3212;
Best Local Similarity 100.0%; Pred. No. 0.0087;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTCGGAAGTGCATCCAAATTCAGGTTTC 28
Db 1192 GTCGGAAGTGCATCCAAATTCAGGTTTC 1219

RESULT 15
AY450925
LOCUS       AY450925               28339 bp      DNA      linear      PRI 04-NOV-2003
DEFINITION   Homo sapiens tumor necrosis factor receptor superfamily, member 6
             (TNFRSF6) gene, complete cds.
ACCESSION    AY450925
VERSION      AY450925.1  GI:38045921
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 28339)
AUTHORS      Livingston, R.J., Rieder, M.J., Chung, M.-W., Ritchie, T.K.,
             Olson, A.N., Nguyen, C.P., Nguyen, D.A., Poel, C.L., Robertson, P.D.,
             Schackwitz, W.S., Sherwood, J.K., Leithauser, B.J. and Nickerson, D.A.
TITLE        Direct Submission
JOURNAL      Submitted (28-OCT-2003) Genome Sciences, University of Washington,
             1705 NE Pacific, Seattle, WA 98195, USA
COMMENT      To cite this work please use: NIEHS-SNPs, Environmental Genome
             Project, NIEHS ES15478, Department of Genome Sciences, Seattle, WA
             (URL: http://egp.gs.washington.edu).
FEATURES     Location/Qualifiers
             1..28339
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
             repeat_region
                 1..263
                     /rpt_family="L2"
                     /rpt_type="dispersed"
             variation
                 310
                     /frequency="0.02"
                     /replace="c"
             variation
                 323
                     /frequency="0.01"
                     /replace="a"
             variation
                 724
                     /frequency="0.09"
                     /replace="c"
             variation
                 744
                     /frequency="0.49"
                     /replace="g"

179
/frequency="0.03"
/replace="a"
1069..26340
/gene="TNFRSF6"
join(1069..1444,13575..13740,18247..18384,19436..19544,
21095..21156,21309..21371,22555..22637,23899..23916,
24668..26340)
/gene="TNFRSF6"
/product="tumor necrosis factor receptor superfamily,
member 6"
1120
/gene="TNFRSF6"
/frequency="0.01"
/replace="a"
1320
/gene="TNFRSF6"
/frequency="0.05"
/replace="a"
1359
/gene="TNFRSF6"
/frequency="0.01"
/replace="a"
join(1415..1444,13575..13740,18247..18384,19436..19544,
21095..21156,21309..21371,22555..22637,23899..23916,
24668..25006)
/gene="TNFRSF6"
/codon_start=1
/product="tumor necrosis factor receptor superfamily,
member 6"
/protein_id="AAR08906.1"
/db_xref="GI:38045922"
/translation="MLGIWTLPLVLTSLVARLSKSNVAQVTIDNSKGLRLKTVTV
ETQNLGHDGQFCHRPFPGERKARDCTVNGDEPCVFCQEGKETDKAHFSSKCR
RCRLDGHGLEVINEICTRQNTKCRCKPNFFCNSTVCEHCDPCTKEGIIKECLT
SNTKCEGSRNLGWLCLLLPLIPLVWVRKEVQTCRKHKENGSGHESPTLNP
LAINFSDVLSKYITTIAGVMTLSQVKGVRKNGVNEAKIDIKNDNQDTAEQKVQ
TLRNWHQLHGKKEAYDTLIKDKKALCTLAEKIQTILKIDTSNSNFRNEIOSL
V"
1479..4261
/gene="TNFRSF6"
/note="Region not scanned for variation"
2462..2686
/rpt_family="MIR"
/rpt_type="dispersed"
4305
/gene="TNFRSF6"
/frequency="0.01"
/replace="a"
4332..4333
/gene="TNFRSF6"
/frequency="0.47"
/replace="a"
4502
/gene="TNFRSF6"
/frequency="0.01"
/replace="g"
4531
/gene="TNFRSF6"
/frequency="0.47"
/replace="c"
4541
/gene="TNFRSF6"
/frequency="0.47"
/replace="g"
4810
/gene="TNFRSF6"
/frequency="0.01"
/replace="c"
4938
/gene="TNFRSF6"
/frequency="0.16"
/replace="c"

```

```

variation
5270 /gene="TNFRSF6"
      /frequency="0.38"
      /replace="c"
5390 /gene="TNFRSF6"
      /frequency="0.47"
      /replace="t"
5527 /gene="TNFRSF6"
      /frequency="0.09"
      /replace="c"
5533 /gene="TNFRSF6"
      /frequency="0.01"
      /replace="g"
5694 /gene="TNFRSF6"
      /frequency="0.48"
      /replace="a"
5738 /gene="TNFRSF6"
      /frequency="0.01"
      /replace="a"
5738 /gene="TNFRSF6"
      /frequency="0.47"
      /replace="t"
5841 /gene="TNFRSF6"
      /frequency="0.47"
      /replace="t"
5977 /gene="TNFRSF6"
      /frequency="0.01"
      /replace="t"
6083 .6392
      /rpt_family="Alu"
      /rpt_type="dispersed"
6143 /gene="TNFRSF6"
      /frequency="0.01"
      /replace="c"
6185 /gene="TNFRSF6"
      /frequency="0.16"
      /replace="t"
6386 /gene="TNFRSF6"
      /frequency="0.75"
      /replace="n"
6596 /gene="TNFRSF6"
      /frequency="0.03"
      /replace="a"
6667 /gene="TNFRSF6"
      /frequency="0.01"
      /replace="c"
6750 /gene="TNFRSF6"
      /frequency="0.49"
      /replace="n"
6821 .6823
      /gene="TNFRSF6"
      /frequency="0.01"
      /replace="n"
6830 /gene="TNFRSF6"
      /frequency="0.50"
      /replace="n"
6853 /gene="TNFRSF6"

```

```

      /frequency="0.49"
      /replace="c"
6884 .7324
      /rpt_family="L1"
      /rpt_type="dispersed"
6934 /gene="TNFRSF6"
      /frequency="0.49"
      /replace="t"
6951 /gene="TNFRSF6"
      /frequency="0.09"
      /replace="c"
6969

Query Match      100.0%; Score 28; DB 9; Length 28339;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GTCGGAACTGCATCCAAATTCAGGTTTC 28
      |||||
Db      380 GTCGGAACTGCATCCAAATTCAGGTTTC 407

Search completed: September 1, 2004, 21:50:45
Job time : 1404 secs

```

This page blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 00:06:29 ; Search time 123 Seconds
(without alignments)
744.271 Million cell updates/sec

Title: US-10-028-415-40

Perfect score: 1747

Sequence: 1 MSSEATQPPAPPAAL.....AADPPAENSSAPAEQGGAE 324

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1747	100.0	324	5 AAU99516	Aau99516 Human YB-
2	1747	100.0	324	6 ABO53035	AbO53035 Human put
3	1747	100.0	369	2 AAR14848	Aar14848 Protein a
4	1747	100.0	369	2 AAW00089	Aaw00089 RAS-relat
5	1747	100.0	369	2 AAY49814	Aay49814 Human gli
6	1747	100.0	369	3 AAB20625	Aab20625 pATG29 hu
7	1743	99.8	384	6 AAO16411	Aao16411 Human nuc
8	1707	97.7	322	2 AAW12827	Aaw12827 Rat TSEP-
9	1707	97.7	322	2 ABB80304	Abb80304 Rat Y-box
10	1707	97.7	322	7 ADE61260	AdE61260 Rat Prote
11	1707	97.7	322	7 ADD48835	AdD48835 Rat Prote
12	1707	97.7	322	7 ADE61263	AdE61263 Rat Prote
13	1701	97.4	322	5 ABB57039	Abb57039 Mouse isc
14	1701	97.4	322	7 ABB80303	Abb80303 Mouse Y-b
15	1699	97.3	322	7 ABB80299	Abb80299 Rat Y-box
16	1699	97.3	322	7 ABB80300	Abb80300 Rat Y-box
17	1699	97.3	322	7 ABB80297	Abb80297 Rat B-TEP
18	1699	96.7	322	7 ABB80302	Abb80302 Mouse Y-b
19	1688	96.6	322	7 ABB80305	Abb80305 Rat Y-box
20	1687	96.6	317	5 AAG70989	Ag70989 Human YB-
21	1687	96.6	317	5 AAU79748	Aau79748 Human DNA
22	1687	96.6	317	7 ABB80298	Abb80298 Human Y-b
23	1687	96.6	318	7 ADD48837	AdD48837 Human Pro
24	1630	93.3	302	4 AAB29744	Aab29744 Human Y-b
25	1534	87.8	290	6 AAE37023	Aae37023 Human nuc

26	1489.5	85.3	375	4 ABG18816	Abg18816 Novel hum
27	865.5	49.5	361	7 ADE56070	AdE56070 Rat Prote
28	840	48.1	372	4 AAU27828	Aau27828 Human ful
29	840	48.1	372	6 ABO53070	AbO53070 Human put
30	840	48.1	442	4 AAC06882	Aac06882 Human pol
31	753.5	43.1	303	4 AAU27829	Aau27829 Dog full-
32	739	42.3	142	4 ABG18815	Abg18815 Novel hum
33	731	41.8	272	3 AAB57119	Aab57119 Human pro
34	657.5	37.6	445	2 AAR14163	Aar14163 Cellular
35	543	31.1	397	4 ABG08725	Abg08725 Novel hum
36	502.5	28.8	121	4 ABG11756	Abg11756 Novel hum
37	500	28.6	315	4 ABB61702	Abb61702 Drosophil
38	497.5	28.5	143	4 ABG11757	Abg11757 Novel hum
39	495	28.3	303	4 AAU18157	Aau18157 Novel hum
40	495	28.3	303	5 ABG92578	Abg92578 Human DNA
41	495	28.3	303	7 ADC25295	AdC25295 Human ext
42	476.5	27.3	253	4 AAM95499	Aam95499 Human rep
43	476.5	27.3	253	4 ABB96182	Abb96182 Human tes
44	476.5	27.3	253	4 AAU18236	Aau18236 Novel hum
45	476.5	27.3	253	5 ABG92657	Abg92657 Human DNA

ALIGNMENTS

RESULT 1

AAU99516

ID AAU99516 standard; protein; 324 AA.

XX AC AAU99516;

DT 07-OCT-2002 (first entry)

XX DE Human YB-1 protein.

XX KW Human; apoptotic cell death; proteinaceous transcription factor;

XX KW regulation of gene transcription; apoptosis; p53; CD95; TRA;

XX KW transcriptional regulator of apoptosis; Y-box family; YB-1; cancer;

XX KW tumour cell; embryonic cell; nervous system; intracellular pathogen;

XX KW DNA-damaging agent; retroviral infection; neurodegenerative disorder;

XX KW immune system dysfunction; anti-tumour; cytostatic.

XX OS Homo sapiens.

XX WO200244363-A1.

XX PD 06-JUN-2002.

XX PF 28-NOV-2001; 2001WO-NZ000287.

XX PR 28-NOV-2000; 2000US-00724809.

XX PA (GENE-) GENESIS RES & DEV CORP LTD.

XX PI Lasham A, Watson JD;

XX PS WPI; 2002-557540/59.

XX DR Modulating p53-mediated apoptotic cell death in a population of cells, by modulating the amount of a transcriptional regulator of apoptosis available to bind to a target polynucleotide in the cells.

XX PT Claim 3; Page 59; 62pp; English.

XX The present invention relates to methods for modulating apoptotic cell death using proteinaceous transcription factors that regulate the transcription of genes encoding proteins involved in apoptosis (e.g. CD95 and p53). The methods involve modulating the amount of a transcriptional regulator of apoptosis (TRA) available to bind to a target polynucleotide in the cells, where TRA is a member of the Y-box nucleic acid binding family of polypeptides (e.g. YB-1). The methods of the invention are useful for modulating apoptotic cell death in a population of cells, where the cells are selected from tumour cells, cells of the immune

AC AAR14848;
XX 25-MAR-2003 (revised)
DT 30-JAN-1992 (first entry)
XX
DE Protein associated with biochemical pathway involving cAMP.
XX
XX RAS; oncogene; cancer; cyclic nucleotide phosphodiesterase.
XX
XX Homo sapiens.
XX
XX WO9116457-A.
XX
XX 31-OCT-1991. 90US-00511715.
XX
XX 20-APR-1990; 90US-00511715.
XX
XX 20-APR-1990; 90US-00511715.
XX
XX (COLD-) COLD SPRING HARBOR LAB.
XX
XX Wigler MH, Colicelli JJ;
XX
XX WPI; 1991-339841/46.
XX N-PSDB; AAQ14635.
XX
XX Complementary screening for genes and prods. - e.g. RAS protein and cAMP,
XX that modify, complement or suppress genetic defect and correct associated
XX phenotypic alteration.
XX
XX Disclosure; Page 114; 169pp; English.
XX
XX In the specification this sequence is given the SEQ ID NO. 31. In the
XX text, SEQ ID NO.31 is described as the nucleotide sequence of the cDNA
XX insert from plasmid pATG29 which was able to complement a mutation in the
XX ras1 gene ! (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 369 AA;
SQ
Query Match 100.0%; Score 1747; DB 2; Length 369;
Best Local Similarity 100.0%; Pred. No. 5.3e-134;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSSEAEQOPPAAPALSAADTKPGTTGSGAGSGGPGGLTSAAPAGDKKVIATKVL 60
Db 46 MSSEAEQOPPAAPALSAADTKPGTTGSGAGSGGPGGLTSAAPAGDKKVIATKVL 105
QY 61 GTVKWFNVRNGYGFINRNDTKEDVFVHQTAKNNPKYLSVGDGETVEFDVVEGEKGA 120
Db 106 GTVKWFNVRNGYGFINRNDTKEDVFVHQTAKNNPKYLSVGDGETVEFDVVEGEKGA 165
QY 121 EAAVTGPGGVPVQGSKYAADRNHYRYPYRRRPPRNYQNYQNSGEGKESGESAPEG 180
Db 166 EAAVTGPGGVPVQGSKYAADRNHYRYPYRRRPPRNYQNYQNSGEGKESGESAPEG 225
QY 181 QAOQRRYRRRPPYPMRPFYGRPOYSNPPVQGEWEGADNQGAGEQGRPVQNNYRG 240
Db 226 QAOQRRYRRRPPYPMRPFYGRPOYSNPPVQGEWEGADNQGAGEQGRPVQNNYRG 285
QY 241 YRPRFRGPRQRPREDGNEEDKENQDGTQGPQRRYRNFYRRRPNPKPDG 300
Db 286 YRPRFRGPRQRPREDGNEEDKENQDGTQGPQRRYRNFYRRRPNPKPDG 345
QY 301 KETKAADPPAENSAPAEQGGAE 324
Db 346 KETKAADPPAENSAPAEQGGAE 369
RESULT 4
AAW00089
ID AAW00089 standard; protein; 369 AA.
XX
AC AAW00089;

XX 25-MAR-2003 (revised)
DT 09-OCT-1996 (first entry)
XX
DE RAS-related protein encoded by plasmid pATG29 (ATCC 68591).
XX
XX Human; glioblastoma; complementation; S. cerevisiae; S. pombe; Clone S46;
KW strain RS60.15B; RAS2; RAS2(vall19ala15); Xenopus laevis;
KW S6 protein kinase; Plasmid pML5; strain SKN37; CAP; adenyllyl cyclase;
KW diploid; strain SP565; ras1; ras1::LEU2; mutation.
XX
XX Homo sapiens.
XX
XX US5527896-A.
XX
XX 18-JUN-1996.
XX
XX 19-APR-1991; 91US-00688352.
XX
XX 20-APR-1990; 90US-00511715.
XX
XX (COLD-) COLD SPRING HARBOR LAB.
XX
XX Colicelli JJ, Wigler MH;
XX
XX WPI; 1996-299902/30.
XX N-PSDB; AAT34371.
XX
XX DNA mols. isolated from human glioblastoma cells - encode RAS-related or
XX cyclic nucleotide phosphodiesterase proteins.
XX
XX Claim 1; Col 107-112; 101pp; English.
XX
XX The sequences given in AAW00087-89 and AAW00101 are encoded by plasmid
XX fragments which contain human glioblastoma cell cDNA inserts which do not
XX encode phosphodiesterases. The cDNA's were obtained by complementation of
XX two genetically altered S. cerevisiae and S. pombe strains. Clone S46 was
XX selected by complementation in S. cerevisiae strain RS60.15B. This strain
XX contains a mutant allele of RAS2, RAS2(vall19ala15), which renders cells
XX unable to grow at 36 deg. C, because such cells are defective in RAS
XX function at elevated temperatures. Human cDNA's from a human glioblastoma
XX cell library were selected that could complement this defect. The deduced
XX sequence of S46 is homologous to a Xenopus laevis gene that encodes a
XX known protein kinase, the S6 protein kinase. Plasmid pML5 was selected by
XX complementation in S. cerevisiae strain SKN37. This strain contains a
XX disrupted allele of CAP, cap::HIS3. CAP encodes an adenyllyl cyclase
XX associated protein of undetermined function. As a consequence of this
XX gene disruption, SKN37 fails to grow in medium rich in amino acids. Human
XX cDNA's were chosen which could complement this defect. Plasmids pATG16
XX and pATG29 were selected by complementation in the S. pombe diploid
XX strain SP565. This strain is homozygous for disruptions of ras1
XX (ras1::LEU2). As a result of this mutation, this strain fails to
XX sporulate, and human cDNA's were selected which could complement this
XX defect. These genes have unknown function. (Updated on 25-MAR-2003 to
XX correct PF field.)
XX
XX Sequence 369 AA;
SQ
Query Match 100.0%; Score 1747; DB 2; Length 369;
Best Local Similarity 100.0%; Pred. No. 5.3e-134;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSSEAEQOPPAAPALSAADTKPGTTGSGAGSGGPGGLTSAAPAGDKKVIATKVL 60
Db 46 MSSEAEQOPPAAPALSAADTKPGTTGSGAGSGGPGGLTSAAPAGDKKVIATKVL 105
QY 61 GTVKWFNVRNGYGFINRNDTKEDVFVHQTAKNNPKYLSVGDGETVEFDVVEGEKGA 120
Db 106 GTVKWFNVRNGYGFINRNDTKEDVFVHQTAKNNPKYLSVGDGETVEFDVVEGEKGA 165
QY 121 EAAVTGPGGVPVQGSKYAADRNHYRYPYRRRPPRNYQNYQNSGEGKESGESAPEG 180
Db 166 EAAVTGPGGVPVQGSKYAADRNHYRYPYRRRPPRNYQNYQNSGEGKESGESAPEG 225
XX

QY 181 QAQORRYRRRRFPYRRPGRPOYSNPPVQGVMEGADNQGAGEQGRPVRRQNNYRG 240
 DB 226 QAQORRYRRRRFPYRRPGRPOYSNPPVQGVMEGADNQGAGEQGRPVRRQNNYRG 285
 QY 241 YRPRFRGPPRQROPREDGNEEDKENQDETQGGQPPQRRYRRNFYRRRRPENPKPDG 300
 DB 286 YRPRFRGPPRQROPREDGNEEDKENQDETQGGQPPQRRYRRNFYRRRRPENPKPDG 345
 QY 301 KETKAADPPAENSAPAEQGGAE 324
 DB 346 KETKAADPPAENSAPAEQGGAE 369
 RESULT 5
 AAY49814
 ID AAY49814 standard; protein; 369 AA.
 XX AC AAY49814;
 XX DT 19-JAN-2000 (first entry)
 XX DE Human glioblastoma cell RAS-related PATG29 protein.
 XX KW Phosphodiesterase; dunce-like phosphodiesterase; PDE; DPD; CAMP;
 XX KW RAS-related protein; immunoreactive; detection; genetic defect;
 XX KW bronchodilation; increased myocardial contractility; anti-inflammation.
 XX OS Homo sapiens.
 XX FN US5977305-A.
 XX PD 02-NOV-1999.
 XX PF 07-JUN-1995; 95US-00474379.
 XX PR 20-APR-1990; 90US-00511715.
 XX PR 19-APR-1991; 91US-00688352.
 XX PR 01-MAR-1994; 94US-00206188.
 XX PA (COLD-) COLD SPRING HARBOR LAB.
 XX PI Colicelli JJ, Wigler MH;
 XX DR WPI; 1999-619709/53.
 XX DR N-PSDB; AAZ32246.
 XX PT New isolated RAS-related polypeptides and mammalian cyclic nucleotide
 phosphodiesterases, used for screening for agents which can modify
 complement or suppress genetic defects.
 XX PS Claim 1; Col 119-122; 145pp; English.
 XX CC The present invention describes new isolated RAS-related polypeptides and
 mammalian cyclic nucleotide phosphodiesterases (PDEs). RAS-related
 polypeptides are capable of complementing a defective RAS function in
 yeast. The products can be used for screening for agents which can
 modify, complement or suppress a genetic defect in a biochemical pathway
 in which CAMP participates, or in a biochemical pathway which is
 controlled, directly or indirectly, by a RAS protein and other proteins
 affecting cell growth and maintenance. Developing agents that will
 selectively act upon PDEs is directed toward reproducing the desirable
 effects of cyclic nucleotides, e.g. bronchodilation, increased myocardial
 contractility, anti-inflammation, yet without causing the undesirable
 effects, e.g. increased heart rate or enhanced lipolysis. The products
 can also be used for therapeutic, diagnostic and prognostic uses.
 CC AAZ32229 to AAZ32285, and AAY49803 to AAY49830, represent sequences used
 in the exemplification of the present invention
 XX SQ Sequence 369 AA;

Query Match 100.0%; Score 1747; DB 2; Length 369;
 Best local Similarity 100.0%; Pred. No. 5.3e-134;

Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSSEAEETQPPAAPAAPALSAADTKPGTTGSGAGSGPGGLTSAAPAGDKKVIATKVL 60
 DB 46 MSSEAEETQPPAAPAAPALSAADTKPGTTGSGAGSGPGGLTSAAPAGDKKVIATKVL 105
 QY 61 GTVKWNVNRYGVGFNRNDTKEDVFNQTAIKKNPKYLRVSGDGTVEFDVVVEGKGA 120
 DB 106 GTVKWNVNRYGVGFNRNDTKEDVFNQTAIKKNPKYLRVSGDGTVEFDVVVEGKGA 165
 QY 121 EAANVTGPGVFGVQSKYAADRNHYRRYRRRPPRPNYQNNYQNSGEKNEGSESAPEG 180
 DB 166 EAANVTGPGVFGVQSKYAADRNHYRRYRRRPPRPNYQNNYQNSGEKNEGSESAPEG 225
 QY 181 QAQORRYRRRRFPYRRPGRPOYSNPPVQGVMEGADNQGAGEQGRPVRRQNNYRG 240
 DB 226 QAQORRYRRRRFPYRRPGRPOYSNPPVQGVMEGADNQGAGEQGRPVRRQNNYRG 285
 QY 241 YRPRFRGPPRQROPREDGNEEDKENQDETQGGQPPQRRYRRNFYRRRRPENPKPDG 300
 DB 286 YRPRFRGPPRQROPREDGNEEDKENQDETQGGQPPQRRYRRNFYRRRRPENPKPDG 345
 QY 301 KETKAADPPAENSAPAEQGGAE 324
 DB 346 KETKAADPPAENSAPAEQGGAE 369
 RESULT 6
 AAB20625
 ID AAB20625 standard; protein; 369 AA.
 XX AC AAB20625;
 XX DT 14-DEC-2000 (first entry)
 XX DE PATG29 human glioblastoma cell insert protein sequence SEQ ID NO:32.
 XX KW Detection; mammalian gene; yeast; microorganism; identification;
 KW KW phenotype; characteristic; dunce-like phosphodiesterase; PDE; RAS;
 KW KW RAS-related protein; genetic defect; hybridisation; probe.
 XX OS Homo sapiens.
 XX OS Schizosaccharomyces pombe.
 XX FN US6100025-A.
 XX PD 08-AUG-2000.
 XX PF 01-MAR-1994; 94US-00206188.
 XX PR 20-APR-1990; 90US-00511715.
 XX PR 19-APR-1991; 91US-00688352.
 XX PA (COLD-) COLD SPRING HARBOR LAB.
 XX PI Colicelli JJ, Wigler MH;
 XX DR WPI; 2000-531664/48.
 XX DR N-PSDB; AAA88181.
 XX PT Novel isolated DNA encoding a mammalian cyclic nucleotide
 phosphodiesterase is present in plasmids pPDE46, pPDE43 or pPDE339 and is
 used to modify a genetic defect in a biochemical pathway in which CAMP
 participates.
 XX PS Example 1; Col 125-128; 145pp; English.
 XX CC The present invention describes a purified and isolated DNA (?) which
 encodes a mammalian cyclic nucleotide phosphodiesterase and is an insert
 present in the plasmids pPDE46 (ATCC 69552), pPDE43 (ATCC 69551) or
 pPDE339 (ATCC 69550). The DNA molecules are used to modify, complement or
 suppress a genetic defect in a biochemical pathway in which CAMP
 participates and are also used as hybridisation probes. The present

CC invention also describes methods for detecting mammalian genes encoding
CC proteins which can function in microorganisms, particularly yeast, to
CC modify, complement, or suppress a genetic defect associated with an
CC identifiable phenotypic alteration or characteristic in the
CC microorganism. AAA88162 to AAA88218 and AAA29614 to AAA29640 represent
CC sequences used in the exemplification of the present invention
XX
SQ Sequence 369 AA;

Query Match 100.0%; Score 1747; DB 3; Length 369;
Best Local Similarity 100.0%; Pred. No. 5.3e-134;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSSEATQPPAAPAAPALSAADTKPGTTGSGAGSGGGLTSAAPAGGKKVIATKVL 60
DB 46 MSSEATQPPAAPAAPALSAADTKPGTTGSGAGSGGGLTSAAPAGGKKVIATKVL 105
QY 61 GTVKWNVNRYGFINRNDTKEDVFVHTAIKKNPKYLSVGDGETVEFDVVEGKGA 120
DB 106 GTVKWNVNRYGFINRNDTKEDVFVHTAIKKNPKYLSVGDGETVEFDVVEGKGA 165
QY 121 EAAVNTGPGVPVQGSKYAADRNHYRYPYRRPQYNNQYNSSEKNGESAPGEG 180
DB 166 EAAVNTGPGVPVQGSKYAADRNHYRYPYRRPQYNNQYNSSEKNGESAPGEG 225
QY 181 QAOQRYPYRRPFPYMYRRPYGRPOYSNPPVQGEVWEGADNOGAGEQGRPVQNNMYRG 240
DB 226 QAOQRYPYRRPFPYMYRRPYGRPOYSNPPVQGEVWEGADNOGAGEQGRPVQNNMYRG 285
QY 241 YRPRFRGPPRQRPREDGNEEDKENQDGTQGGQPPQRRYRNFNRYRRRPNPKPDG 300
DB 286 YRPRFRGPPRQRPREDGNEEDKENQDGTQGGQPPQRRYRNFNRYRRRPNPKPDG 345
QY 301 KETKAADPPAENSAPAEQGGAE 324
DB 346 KETKAADPPAENSAPAEQGGAE 369

RESULT 7
AA016411
ID AA016411 standard; protein; 384 AA.
AC AA016411;
XX
XX
DT 10-APR-2003 (first entry)
XX
DE Human nucleic acid-associated protein (NAAP) - SEQ ID No 8.
XX
XX Human; nucleic acid-associated protein; NAAP; arteriosclerosis;
KW cell proliferative disorder; atherosclerosis; cirrhosis; hepatitis; AIDS;
KW cancer; developmental disorder; renal tubular acidosis; anaemia; asthma;
KW mental retardation; neurological disorder; Alzheimer's disease; epilepsy;
KW Parkinson's disease; autoimmune disorder; inflammatory disorder; allergy;
KW Crohn's disease; transgenic animal; animal model.
XX
OS Homo sapiens.
XX
XX WO2003000864-A2.
XX
XX PD 03-JAN-2003.
XX
XX PF 20-JUN-2002; 2002WO-US021179.
XX
XX 22-JUN-2001; 2001US-0300518P.
XX PR 29-JUN-2001; 2001US-0301787P.
XX PR 29-JUN-2001; 2001US-0301792P.
XX PR 29-JUN-2001; 2001US-0301892P.
XX PR 29-JUN-2001; 2001US-0301893P.
XX PR 06-JUL-2001; 2001US-0303405P.
XX PR 06-JUL-2001; 2001US-0303442P.
XX PR 15-MAR-2002; 2002US-0364438P.
XX
XX PA (INCY-) INCYTE GENOMICS INC.

XX Gandhi AR, Swarnakar A, Hafalia AJA, Warren BA, Emerling BM;
PI Arvizu CS, Ison CH, Honchell CD, Lee EA, Yue H, Forsythe IO;
PI Ramkumar J, Griffin JA, Yang J, Sanjanwala MM, Baughn MR;
PI Borowsky MJ, Yao MG, Walla NK, Bandman O, Lal PG, Becha SD, Lee SY;
PI Richardson TW, Elliott VS, Luo W, Tang VT, Zebardjian Y, Lu Y;
XX
DR WPI: 2003-201420/19.
DR N-PSDB; AAL51561.
XX
PT New nucleic acid-associated proteins and polynucleotides, useful for
PT diagnosing, treating or preventing cell proliferative (e.g. cancer),
PT neurological (e.g. epilepsy or Parkinson's disease), or autoimmune
PT disorders (e.g. AIDS).
XX
PS Claim 1; Page 222-223; 312pp; English.
XX
CC The invention comprises the amino acid and coding sequences of human
CC nucleic acid-associated proteins (NAAP). The DNA and protein sequences of
CC the invention are useful for diagnosing, treating or preventing disorders
CC associated with aberrant expression of NAAP, such as: cell proliferative
CC disorders (e.g. arteriosclerosis, atherosclerosis, cirrhosis, hepatitis,
CC or cancer); developmental disorders (e.g. renal tubular acidosis, anaemia
CC or mental retardation); neurological disorders (e.g. Alzheimer's disease,
CC Parkinson's disease or epilepsy); and autoimmune/inflammatory disorders
CC (e.g. AIDS, allergies, asthma or Crohn's disease). The DNA sequences of
CC the invention are useful for creating transgenic animals to model human
CC disease. The present amino acid sequence represents a human nucleic acid-
CC associated protein of the invention
XX
SQ Sequence 384 AA;

Query Match 99.8%; Score 1743; DB 6; Length 384;
Best Local Similarity 99.7%; Pred. No. 1.2e-133;
Matches 323; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSSEATQPPAAPAAPALSAADTKPGTTGSGAGSGGGLTSAAPAGGKKVIATKVL 60
DB 61 MSSEATQPPAAPAAPALSAADTKPGTTGSGAGSGGGLTSAAPAGGKKVIATKVL 120
QY 61 GTVKWNVNRYGFINRNDTKEDVFVHTAIKKNPKYLSVGDGETVEFDVVEGKGA 120
DB 121 GTVKWNVNRYGFINRNDTKEDVFVHTAIKKNPKYLSVGDGETVEFDVVEGKGA 180
QY 121 EAAVNTGPGVPVQGSKYAADRNHYRYPYRRPQYNNQYNSSEKNGESAPGEG 180
DB 181 EAAVNTGPGVPVQGSKYAADRNHYRYPYRRPQYNNQYNSSEKNGESAPGEG 240
QY 181 QAOQRYPYRRPFPYMYRRPYGRPOYSNPPVQGEVWEGADNOGAGEQGRPVQNNMYRG 240
DB 241 QAOQRYPYRRPFPYMYRRPYGRPOYSNPPVQGEVWEGADNOGAGEQGRPVQNNMYRG 300
QY 241 YRPRFRGPPRQRPREDGNEEDKENQDGTQGGQPPQRRYRNFNRYRRRPNPKPDG 300
DB 301 YRPRFRGPPRQRPREDGNEEDKENQDGTQGGQPPQRRYRNFNRYRRRPNPKPDG 360
QY 301 KETKAADPPAENSAPAEQGGAE 324
DB 361 KETKAADPPAENSAPAEQGGAE 384

RESULT 8
AA012827
ID AA012827 standard; protein; 322 AA.
XX
XX AC AA012827;
XX
DT 08-JAN-1998 (first entry)
XX
XX DE Rat TSEP-1 protein.
XX
XX Promoter; PD1; major histocompatibility complex; MHC Class I molecule;
KW mammalian cell; regulator; autoimmune disease; transplant rejection;
KW

therapy; non-thyroid autoimmune disease; Sox-4; TSEP-1 protein.

Rattus rattus.

Key Location/Qualifiers

FT Misc-difference 148..150 /note= "possible nuclear localisation signal"

FT Misc-difference 183..190 /note= "possible nuclear localisation signal"

FT Misc-difference 197..203 /note= "possible nuclear localisation signal"

FT Misc-difference 240..245 /note= "possible nuclear localisation signal"

FT Misc-difference 277..281 /note= "possible nuclear localisation signal"

FT Misc-difference 286..289 /note= "possible nuclear localisation signal"

WO9707404-A1.

27-FEB-1997.

21-AUG-1996; 96WO-US013715.

21-AUG-1995; 95US-00503525.

(USSH) US SEC DEPT HEALTH.

Kohn L, Singer DS, Saji M, Giuliani C, Shong M, Suzuki K, Ohmori M;

WPI; 1997-165448/15.

N-PSDB; AAT59845.

Identification of drugs which modulate MHC Class I expression - used in the prevention and treatment of autoimmune disease and transplant rejection.

Claim 23; Fig 38b; 273pp; English.

This sequence represents the rat TSEP-1 protein. This protein is a Y-box protein, and can be used in the method of the invention. The method of the invention is for assessing the ability, or therapeutic potential, of a drug to suppress expression of MHC Class I molecules by measuring altered binding of a mammalian cell protein extract to an MHC Class I regulatory nucleic acid sequence or functional equivalent. The method comprises treating the mammalian cells with the drug. Protein extract is then obtained from the cells and is combining with a regulatory nucleic acid sequence to allow formation of at least one complex between the protein and nucleic acid. Finally, the formation of the complex is detected where altered complex formation indicates the drugs potential in treating autoimmune disease or transplant rejection. The other methods of the invention are for assessing the ability of a drug to suppress expression of MHC Class I molecules, for assessing the therapeutic potential of a candidate drug, and for assessing the therapeutic potential of a candidate drug for treating autoimmune disease or transplant rejection. The methods are used to identify nucleic acids and other compounds capable of MHC Class I suppression which may be used in the prevention or treatment of transplant rejection in a mammal, or the treatment of a non-thyroid autoimmune disease. The antibodies of the invention may be used to detect either Sox-4 or TSEP-1 proteins

Sequence 322 AA;

Query Match 97.7%; Score 1707; DB 2; Length 322;

Best Local Similarity 98.5%; Pred. No. 8.2e-131; Mismatches 2; Gaps 1; Matches 319; Conservative 1;

QY 1 MSSEAEQQPPAAPALSAADTKPGTTGSGAGSGPGGLTSAAPAGGDKKVIATKYL 60
 Db 1 MSSEAEQQPPAAP--AAALSAADTKPGSTGSGAGSGPGGLTSAAPAGGDKKVIATKYL 58

QY 61 GTVKWFWNRNGYGFINRNDTKEDVFVHQTAKKNPKYLSVGDGETVEFDVVEGEKGA 120

Db 59 GTVKWFWNRNGYGFINRNDTKEDVFVHQTAKKNPKYLSVGDGETVEFDVVEGEKGA 118

QY 121 EAANTVGGGVPVQGSKYAADRNHYRRYPRRRGPPRYQNYQNSGEKNGESASEG 180

Db 119 EAANTVGGGVPVQGSKYAADRNHYRRYPRRRGPPRYQNYQNSGEKNGESASEG 178

QY 181 QAQQRPPYRRRRFPYNNRRYPGRPPQYSNPPVQGVNVEGADNCGAGQGPRVQNNMYRG 240

Db 179 QAQQRPPYRRRRFPYNNRRYPGRPPQYSNPPVQGVNVEGADNCGAGQGPRVQNNMYRG 238

QY 241 YRPRFRGPPRQRPREDGNEEDKENQGDETQGGQPPORRYRRNFRNRRRPNPKPDG 300

Db 239 YRPRFRGPPRQRPREDGNEEDKENQGDETQGGQPPORRYRRNFRNRRRPNPKPDG 298

QY 301 KETKAADPPAENSSAPEAEQGGAE 324

Db 299 KETKAADPPAENSSAPEAEQGGAE 322

RESULT 9

ABB80304

ID ABB80304 standard; protein; 322 AA.

XX AC ABB80304;

XX DT 15-JAN-2004 (first entry)

XX DE Rat Y-box protein.

XX KW Protein tyrosine phosphatase 1B; PTP1B; insulin receptor kinase; IRK;
 protein tyrosine kinase; diabetes; obesity; Y-box protein; YB-1;
 impaired glucose tolerance.

XX OS Rattus norvegicus.

XX PN WO2003076634-A2.

XX PD 18-SEP-2003.

XX PF 12-MAR-2003; 2003WO-EP002552.

XX PR 12-MAR-2002; 2002US-0363787P.

XX PR 20-DEC-2002; 2002US-0435587P.

XX PR 11-MAR-2003; 2003US-00388215.

XX PA (COLD-) COLD SPRING HARBOR LAB.

XX PI (TONK/) TONKS N K.

XX PI Fukada T;

XX DR WPI; 2003-756826/71.

XX PT New isolated polynucleotide useful for treating protein tyrosine phosphatase (PTP)1B-associated disorders (e.g. diabetes, obesity or impaired glucose tolerance) and for identifying agents capable of altering PTP1B expression.

XX PS Claim 19; Page 113-14; 115pp; English.

XX CC This sequence represents a rat Y-box protein. This protein binds to the Y-box protein (YB-1) binding site. Y-box proteins are capable of interacting with DNA and RNA and possess a conserved domain of approx. 70 amino acids termed the cold-shock domain, which is capable of binding to a DNA "Y-box" sequence (ATTGG). This sequence is useful in treating PTP1B-associated disorders or conditions associated with defects in biological signal transduction in cells, such as diabetes, obesity, impaired glucose tolerance and other metabolic disorders. This sequence may also be used in identifying agents capable of altering PTP1B expression

XX Sequence 322 AA;

Query Match 97.7%; Score 1707; DB 7; Length 322;

Best Local Similarity 98.5%; Pred. No. 8.2e-131;
Matches 319; Conservative 1; Mismatches 2; Indels 2; Gaps 1;
QY 1 MSSEAEATQPPAAPALSAADTKPGTTGSGAGSGPGGLTSAAPAGGDKKVIATKVL 60
DB 1 MSSEAEATQPPAAP--AAALSAADTKPGTSGAGSGPGGLTSAAPAGGDKKVIATKVL 58
QY 61 GTVKWNVNRYGFIINRNDTKEDVHVQTAIKKNPKYLSVGDGETVEFDVVEGEKGA 120
DB 59 GTVKWNVNRYGFIINRNDTKEDVHVQTAIKKNPKYLSVGDGETVEFDVVEGEKGA 118
QY 121 EAAVNTGPGGVPVQGSKYAADRNHYRYPRRRPPPNYQONYSSEGEKNGESAPRG 180
DB 119 EAAVNTGPGGVPVQGSKYAADRNHYRYPRRRPPPNYQONYSSEGEKNGESAPRG 178
QY 181 QAQORRPPYRRRPPYMYRRPYGRPPQYNSNPPVQGEVMEGADNQGAGEQGRPVQNMRYG 240
DB 179 QAQORRPPYRRRPPYMYRRPYGRPPQYNSNPPVQGEVMEGADNQGAGEQGRPVQNMRYG 238
QY 241 YRFRFRGPPRQPRQPRDGNEDKENGQDETQCGQPPORRYRNFNRYRRRPNPKPDG 300
DB 239 YRFRFRGPPRQPRQPRDGNEDKENGQDETQCGQPPORRYRNFNRYRRRPNPKPDG 298
QY 301 KETKAADPPAENSAPAEQGGAE 324
DB 299 KETKAADPPAENSAPAEQGGAE 322

RESULT 10
ADE61260
ID ADE61260 standard; protein; 322 AA.

AC ADE61260;
XX
DT 29-JAN-2004 (first entry)

XX Rat Protein: P27817, SEQ ID NO 7178.
DE Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.
XX WO2003016475-A2.
XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.
XX 14-AUG-2001; 2001US-0312147P.
XX 01-NOV-2001; 2001US-0346382P.
XX 26-NOV-2001; 2001US-0333347P.

XX (GEO) GEN HOSPITAL CORP.
XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
XX GENBANK; P27817.

XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent

CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 322 AA;

Query Match 97.7%; Score 1707; DB 7; Length 322;
Best Local Similarity 98.5%; Pred. No. 8.2e-131;
Matches 319; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 1 MSSEAEATQPPAAPALSAADTKPGTTGSGAGSGPGGLTSAAPAGGDKKVIATKVL 60
DB 1 MSSEAEATQPPAAP--AAALSAADTKPGTSGAGSGPGGLTSAAPAGGDKKVIATKVL 58

QY 61 GTVKWNVNRYGFIINRNDTKEDVHVQTAIKKNPKYLSVGDGETVEFDVVEGEKGA 120
DB 59 GTVKWNVNRYGFIINRNDTKEDVHVQTAIKKNPKYLSVGDGETVEFDVVEGEKGA 118

QY 121 EAAVNTGPGGVPVQGSKYAADRNHYRYPRRRPPPNYQONYSSEGEKNGESAPRG 180
DB 119 EAAVNTGPGGVPVQGSKYAADRNHYRYPRRRPPPNYQONYSSEGEKNGESAPRG 178

QY 181 QAQORRPPYRRRPPYMYRRPYGRPPQYNSNPPVQGEVMEGADNQGAGEQGRPVQNMRYG 240
DB 179 QAQORRPPYRRRPPYMYRRPYGRPPQYNSNPPVQGEVMEGADNQGAGEQGRPVQNMRYG 238

QY 241 YRFRFRGPPRQPRQPRDGNEDKENGQDETQCGQPPORRYRNFNRYRRRPNPKPDG 300
DB 239 YRFRFRGPPRQPRQPRDGNEDKENGQDETQCGQPPORRYRNFNRYRRRPNPKPDG 298

QY 301 KETKAADPPAENSAPAEQGGAE 324

DB 299 KETKAADPPAENSAPAEQGGAE 322

RESULT 11

ADD48835
ID ADD48835 standard; protein; 322 AA.

XX AC ADD48835;
XX 29-JAN-2004 (first entry)

XX Rat Protein A23677, SEQ ID NO 14545.
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.
XX WO2003016475-A2.

XX 27-FEB-2003.
XX 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX (GEO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX Woolf C, D'urso D, Befort K, Costigan M;
 XX WPI; 2003-268312/26.
 DR GENBANK; A23677.
 XX
 XX New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 PS Claim 1; Page; 1017pp; English.
 XX
 XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (SNI)), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 322 AA;
 Query Match 97.7%; Score 1707; DB 7; Length 322;
 Best Local Similarity 98.5%; Pred. No. 8.2e-131;
 Matches 319; Conservative 1; Mismatches 2; Indels 2; Gaps 1;
 QY 1 MSSEAFQCPAPPAALPALSADTKPTGSGAGSGGPGGLTSAPAGDKKVIATKYL 60
 DB 1 MSSEAFQCPAPPAALPALSADTKPTGSGAGSGGPGGLTSAPAGDKKVIATKYL 58
 QY 61 GTVKWFVNRNGYGFINRDKEDVFVHQTAIKNNPKVLSRSGDGETVEFDVVEGEKA 120
 DB 59 GTVKWFVNRNGYGFINRDKEDVFVHQTAIKNNPKVLSRSGDGETVEFDVVEGEKA 118
 QY 121 EAAVNTGPGVPGVQSGKYADRHHYRYPRRRPPRYQNYQNSGSGKNGESAPGG 180
 DB 119 EAAVNTGPGVPGVQSGKYADRHHYRYPRRRPPRYQNYQNSGSGKNGESAPGG 178
 QY 181 QAQORRRYRRPPPPYMYRRPQYRPPQYSNPPVQGVMEGADNQGAGEQGRPVQRMVYRG 240
 DB 179 QAQORRRYRRPPPPYMYRRPQYRPPQYSNPPVQGVMEGADNQGAGEQGRPVQRMVYRG 238
 QY 241 YRFRFRGPPRQRPREDGNEEDKENQGBETQCCQPPQRRYRNFYRRRRRPNPKPDG 300
 DB 239 YRFRFRGPPRQRPREDGNEEDKENQGBETQCCQPPQRRYRNFYRRRRRPNPKPDG 298
 QY 301 KETKAADPPAENSSAPEAEOGGAE 324

DB 299 KETKAADPPAENSSAPEAEOGGAE 322
 RESULT 12
 ADE61263
 ID ADE61263 standard; protein; 322 AA.
 XX
 AC ADE61263;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Rat Protein P27817, SEQ ID NO 7181.
 XX
 KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
 XX
 OS Rattus norvegicus.
 XX
 PN WO2003016475-A2.
 XX
 PD 27-FEB-2003.
 XX
 PF 14-AUG-2002; 2002WO-US025765.
 XX
 PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 PA (GEO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX
 XX Woolf C, D'urso D, Befort K, Costigan M;
 PI WPI; 2003-268312/26.
 DR GENBANK; P27817.
 XX
 XX New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 PT Claim 1; Page; 1017pp; English.
 XX
 XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (SNI)), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 322 AA;
 Query Match 97.7%; Score 1707; DB 7; Length 322;
 Best Local Similarity 98.5%; Pred. No. 8.2e-131;
 Matches 319; Conservative 1; Mismatches 2; Indels 2; Gaps 1;
 QY 1 MSSEAFQCPAPPAALPALSADTKPTGSGAGSGGPGGLTSAPAGDKKVIATKYL 60
 DB 1 MSSEAFQCPAPPAALPALSADTKPTGSGAGSGGPGGLTSAPAGDKKVIATKYL 58
 QY 61 GTVKWFVNRNGYGFINRDKEDVFVHQTAIKNNPKVLSRSGDGETVEFDVVEGEKA 120
 DB 59 GTVKWFVNRNGYGFINRDKEDVFVHQTAIKNNPKVLSRSGDGETVEFDVVEGEKA 118
 QY 121 EAAVNTGPGVPGVQSGKYADRHHYRYPRRRPPRYQNYQNSGSGKNGESAPGG 180
 DB 119 EAAVNTGPGVPGVQSGKYADRHHYRYPRRRPPRYQNYQNSGSGKNGESAPGG 178
 QY 181 QAQORRRYRRPPPPYMYRRPQYRPPQYSNPPVQGVMEGADNQGAGEQGRPVQRMVYRG 240
 DB 179 QAQORRRYRRPPPPYMYRRPQYRPPQYSNPPVQGVMEGADNQGAGEQGRPVQRMVYRG 238
 QY 241 YRFRFRGPPRQRPREDGNEEDKENQGBETQCCQPPQRRYRNFYRRRRRPNPKPDG 300
 DB 239 YRFRFRGPPRQRPREDGNEEDKENQGBETQCCQPPQRRYRNFYRRRRRPNPKPDG 298
 QY 301 KETKAADPPAENSSAPEAEOGGAE 324

Matches 319; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 1 MSSEAEATQPPAAPALSAADTKPGTTGSGAGSGGGLTSAAPAGGDKKVIATKVL 60
 DB 1 MSSEAEATQPPAAP--AAALSAADTKPGTSGAGSGGGLTSAAPAGGDKKVIATKVL 58
 QY 61 GTVKWFNVNRYGFIINRNDTKEDVFVHQTAKKNNPRKYLRSVGDGETVEFDVVEGEKGA 120
 DB 59 GTVKWFNVNRYGFIINRNDTKEDVFVHQTAKKNNPRKYLRSVGDGETVEFDVVEGEKGA 118
 QY 121 EAAANTGCGGVPVQGSKYAADRNHYRYPYRRGPPPNYQNYQNSSEKNEGSESAPSG 180
 DB 119 EAAANTGCGGVPVQGSKYAADRNHYRYPYRRGPPPNYQNYQNSSEKNEGSESAPSG 178
 QY 181 QAQORRPFYRRRRFPYMYMRFPYGRPOYSNPPVQGEVMEGADNOGAGGQGRPVQNNMYRG 240
 DB 179 QAQORRPFYRRRRFPYMYMRFPYGRPOYSNPPVQGEVMEGADNOGAGGQGRPVQNNMYRG 238
 QY 241 YRFRFRGPPRQRPREDGNEEDKENQDGTQCGQPPORRYRRNFYRRRPNPKPDQG 300
 DB 239 YRFRFRGPPRQRPREDGNEEDKENQDGTQCGQPPORRYRRNFYRRRPNPKPDQG 298
 QY 301 KETKAADPPAENSAPAEQGGAE 324
 DB 299 KETKAADPPAENSAPAEQGGAE 322

RESULT 13
 ABB57039
 ID ABB57039 standard; protein; 322 AA.
 XX
 AC ABB57039;
 DT 07-MAR-2002 (first entry)
 XX
 DE Mouse ischaemic condition related protein sequence SEQ ID NO:52.
 XX
 KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
 KW vasospastic ischaemia; ischaemic condition; ischaemic disease.
 XX
 OS Mus musculus.
 XX
 FN WO200188188-A2.
 XX
 PD 22-NOV-2001.
 XX
 PF 18-MAY-2001; 2001WO-JP004192.
 XX
 PR 18-MAY-2000; 2000JP-00145977.
 XX
 PA (UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.
 XX
 PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
 XX
 DR WPI; 2002-034733/04.
 XX
 DR N-PSDB; ABI99233.
 XX
 PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
 PT expression levels of particular genes defined in the specification or by
 PT determining the expression profile of a gene group comprising these
 PT genes.
 XX
 PS Claim 2; Page 171-172; 2690pp; English.
 XX
 CC The present invention describes a method for examining ischaemic
 CC conditions, comprising measuring the expression levels of particular
 CC genes (i) in a test sample or determining the expression profile of a
 CC gene group in the sample comprising genes selected from (i). The method
 CC is useful for examining the ischaemic condition (e.g. compressive
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
 CC expression levels of particular genes (ABI99202 to ABI99912, encoding the
 CC protein sequences in ABB57020 to ABB57374) or by determining the
 CC expression profile of a gene group comprising these genes. The expression

CC levels or expression profiles produced by these genes are used as an
 CC indicator when screening for ischaemic condition-improving drugs or
 CC therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR
 CC primers for a mouse ischaemic condition related sequence, which are used
 CC in the exemplification of the present invention
 XX
 SQ Sequence 322 AA;

Query Match 97.4%; Score 1701; DB 5; Length 322;
 Best Local Similarity 98.1%; Pred. No. 2.5e-130;
 Matches 318; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 1 MSSEAEATQPPAAPALSAADTKPGTTGSGAGSGGGLTSAAPAGGDKKVIATKVL 60
 DB 1 MSSEAEATQPPAAP--AAALSAADTKPGTSGAGSGGGLTSAAPAGGDKKVIATKVL 58
 QY 61 GTVKWFNVNRYGFIINRNDTKEDVFVHQTAKKNNPRKYLRSVGDGETVEFDVVEGEKGA 120
 DB 59 GTVKWFNVNRYGFIINRNDTKEDVFVHQTAKKNNPRKYLRSVGDGETVEFDVVEGEKGA 118
 QY 121 EAAANTGCGGVPVQGSKYAADRNHYRYPYRRGPPPNYQNYQNSSEKNEGSESAPSG 180
 DB 119 EAAANTGCGGVPVQGSKYAADRNHYRYPYRRGPPPNYQNYQNSSEKNEGSESAPSG 178
 QY 181 QAQORRPFYRRRRFPYMYMRFPYGRPOYSNPPVQGEVMEGADNOGAGGQGRPVQNNMYRG 240
 DB 179 QAQORRPFYRRRRFPYMYMRFPYGRPOYSNPPVQGEVMEGADNOGAGGQGRPVQNNMYRG 238
 QY 241 YRFRFRGPPRQRPREDGNEEDKENQDGTQCGQPPORRYRRNFYRRRPNPKPDQG 300
 DB 239 YRFRFRGPPRQRPREDGNEEDKENQDGTQCGQPPORRYRRNFYRRRPNPKPDQG 298
 QY 301 KETKAADPPAENSAPAEQGGAE 324
 DB 299 KETKAADPPAENSAPAEQGGAE 322

RESULT 14
 ABB80303
 ID ABB80303 standard; protein; 322 AA.
 XX
 AC ABB80303;
 DT 15-JAN-2004 (first entry)
 XX
 DE Mouse Y-box protein.
 XX
 KW Protein tyrosine phosphatase 1B; PTP1B; insulin receptor kinase; IRK;
 KW protein tyrosine kinase; diabetes; obesity; Y-box protein; YB-1;
 KW impaired glucose tolerance.
 XX
 OS Mus musculus.
 XX
 FN WO2003076634-A2.
 XX
 PD 18-SEP-2003.
 XX
 PF 12-MAR-2003; 2003WO-BP002552.
 XX
 PR 12-MAR-2002; 2002US-0363787P.
 XX
 PR 20-DEC-2002; 2002US-0435587P.
 XX
 PR 11-MAR-2003; 2003US-00388215.
 XX
 PA (COLD-) COLD SPRING HARBOR LAB.
 XX
 PI (TONK/) TONKS N K.
 XX
 PI Fukada T;
 DR WPI; 2003-756826/71.
 XX
 PT New isolated polynucleotide useful for treating protein tyrosine
 PT phosphatase (PTP)1B-associated disorders (e.g. diabetes, obesity or
 PT impaired glucose tolerance) and for identifying agents capable of

PT altering FPIB expression.
 PS Claim 19; Page 112-13; 115pp; English.
 XX
 CC This sequence represents a mouse Y-box protein. This protein binds to the
 CC Y-box protein (YB-1) binding site. Y-box proteins are capable of
 CC interacting with DNA and RNA and possess a conserved domain of approx. 70
 CC amino acids termed the cold-shock domain, which is capable of binding to
 CC a DNA "Y-box" sequence (ATGGG). This sequence is useful in treating FPIB
 CC -associated disorders or conditions associated with defects in biological
 CC signal transduction in cells, such as diabetes, obesity, impaired glucose
 CC tolerance and other metabolic disorders. This sequence may also be used
 CC in identifying agents capable of altering FPIB expression
 XX
 SQ Sequence 322 AA;
 Query Match 97.4%; Score 1701; DB 7; Length 322;
 Best Local Similarity 98.1%; Pred. No. 2.5e-130;
 Matches 318; Conservative 1; Mismatches 3; Indels 2; Gaps 1;
 QY 1 MSSEAEIQPPAAPPALSAADTKPGTTGSGAGSGGPGGLTSAAPAGGDKKVIATKVL 60
 DB 1 MSSEAEIQPPAAPPALSAADTKPGTTGSGAGSGGPGGLTSAAPAGGDKKVIATKVL 58
 QY 61 GTVKWFNVRNGYGFINRNDTKEDVFVHQTAKKNNPKYLSVGDGETVEFDVVEGEKGA 120
 DB 59 GTVKWFNVRNGYGFINRNDTKEDVFVHQTAKKNNPKYLSVGDGETVEFDVVEGEKGA 118
 QY 121 EAAVNTGPGGVVQGSKYAADRNHYRYPYRRGPPRYQYQNSSEKNEGESAPEG 180
 DB 119 EAAVNTGPGGVVQGSKYAADRNHYRYPYRRGPPRYQYQNSSEKNEGESAPEG 178
 QY 181 QAOQRYPYRRRFPYMYRRPYGRRPOYSNPPVQGEVMEGADNQAGEQGRPVQRNNYRG 240
 DB 179 QAOQRYPYRRRFPYMYRRPYGRRPOYSNPPVQGEVMEGADNQAGEQGRPVQRNNYRG 238
 QY 241 YRPRFRGPPRQRPREDGNEEDKENQDGTQGGQPPQRRYRNFYRRRRRPNPKPDG 300
 DB 239 YRPRFRGPPRQRPREDGNEEDKENQDGTQGGQPPQRRYRNFYRRRRRPNPKPDG 298
 QY 301 KETKAADPPAENSAPAEQGGAE 324
 DB 299 KETKAADPPAENSAPAEQGGAE 322

Search completed: August 25, 2004, 00:23:47
 Job time : 127 secs

RESULT 15
 ABB80299
 ID ABB80299 standard; protein; 322 AA.

XX ABB80299;
 XX AC
 XX 15-JAN-2004 (first entry)

DE Rat Y-box protein.

XX Protein tyrosine phosphatase 1B; FPIB; insulin receptor kinase; IRK;
 XX protein tyrosine kinase; diabetes; obesity; Y-box protein; YB-1;
 XX impaired glucose tolerance.

OS Rattus norvegicus.

XX WO2003076634-A2.

XX 18-SEP-2003.

XX 12-MAR-2003; 2003WO-BP002552.

XX 12-MAR-2002; 2002US-0363787P.

XX 20-DEC-2002; 2002US-0435587P.

XX 11-MAR-2003; 2003US-00388215.

XX (COLD-) COLD SPRING HARBOR LAB.
 PA (TUNK/) TONKS N K.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 00:21:45 ; Search time 32 Seconds
(without alignments)
522.713 Million cell updates/sec

Title: US-10-028-415-40

Perfect score: 1747

Sequence: 1 MSSEAEQQPPAAPAPAL.....AADPPAENGSAPEQQGAE 324

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1747	100.0	369	1	US-07-688-352C-32
2	1747	100.0	369	2	US-08-474-379C-32
3	1747	100.0	369	3	US-09-146-249A-32
4	1747	100.0	369	3	US-08-206-158B-32
5	1747	100.0	369	5	PCT-US91-02714-31
6	384	22.0	74	1	US-08-203-806B-13
7	384	22.0	74	4	US-09-017-754A-13
8	172	9.8	131	4	US-09-489-039A-13476
9	168.5	9.6	82	4	US-09-134-001C-5476
10	164.5	9.4	116	4	US-09-252-991A-19664
11	163.5	9.4	86	4	US-09-107-532A-5955
12	163	9.3	69	1	US-08-203-806B-4
13	163	9.3	69	1	US-08-203-806B-9
14	163	9.3	69	4	US-09-017-754A-4
15	163	9.3	69	4	US-09-017-754A-9
16	163	9.3	73	4	US-09-489-039A-14126
17	163	9.3	507	4	US-09-252-991A-17308
18	162.5	9.3	214	4	US-09-252-991A-29350
19	162	9.3	70	4	US-08-516-667-80
20	160.5	9.2	94	4	US-09-540-236-2432
21	159.5	9.1	79	4	US-09-134-001C-3081
22	158	9.0	73	4	US-09-543-681A-4308
23	158	9.0	99	4	US-09-543-681A-5351
24	156.5	9.0	481	4	US-09-252-991A-17707
25	155	8.9	82	4	US-09-543-681A-7943
26	152.5	8.7	1075	4	US-09-252-991A-18387
27	152	8.7	71	1	US-08-203-806B-2

28	152	8.7	71	1	US-08-203-806B-8	Sequence 8, Appl
29	152	8.7	71	4	US-09-017-754A-2	Sequence 2, Appl
30	152	8.7	71	4	US-09-017-754A-8	Sequence 8, Appl
31	152	8.7	71	4	US-09-516-667-81	Sequence 81, Appl
32	152	8.7	71	4	US-09-293-427-4	Sequence 4, Appl
33	151.5	8.7	71	4	US-09-328-352-4376	Sequence 4376, Ap
34	151.5	8.7	73	4	US-09-107-532A-6235	Sequence 6235, Ap
35	150.5	8.6	352	4	US-09-252-991A-24482	Sequence 24482, A
36	150.5	8.6	705	4	US-09-252-991A-30792	Sequence 30792, A
37	150	8.6	70	1	US-08-203-806B-7	Sequence 7, Appl
38	150	8.6	70	4	US-09-017-754A-7	Sequence 7, Appl
39	150	8.6	70	4	US-09-017-754A-20	Sequence 20, Appl
40	150	8.6	70	4	US-09-516-667-82	Sequence 82, Appl
41	150	8.6	72	4	US-09-489-039A-10823	Sequence 10823, A
42	150	8.6	1098	4	US-09-252-991A-20186	Sequence 20186, A
43	149	8.5	85	4	US-09-540-236-2549	Sequence 2549, Ap
44	148	8.5	720	4	US-09-252-991A-21881	Sequence 21881, A
45	147.5	8.4	396	4	US-09-252-991A-24697	Sequence 24697, A

ALIGNMENTS

RESULT 1
US-07-688-352C-32
; Sequence 32, Application US/07688352C
; Patent No. 5527896
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: Cloning by Complementation and Related
; TITLE OF INVENTION: Processes
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07688,352C
; FILING DATE: 19910419
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,715
; FILING DATE: 20-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25447
; REFERENCE/DOCKET NUMBER: 27805/30197
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 369 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-688-352C-32

Query Match 100.0%; Score 1747; DB 1; Length 369;
Best Local Similarity 100.0%; Pred. No. 1.1e-139;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSEAEETQPPAAPALSAADTKGTTGSGAGSGGPGGLTSAAPAGGDKKVIATKVL 60
 Db 46 MSSEAEETQPPAAPALSAADTKGTTGSGAGSGGPGGLTSAAPAGGDKKVIATKVL 105
 QY 61 GTVKFNVNRYGGINRNDTKEDVVFHQTAKONPKYLSVGDGTVEFDVVEGEKGA 120
 Db 106 GTVKFNVNRYGGINRNDTKEDVVFHQTAKONPKYLSVGDGTVEFDVVEGEKGA 165
 QY 121 EAAVTGPGGVVQGSKYAADRNHYRYPYRRGPPRYNYQNSSEGEKNEGSSESAPEG 180
 Db 166 EAAVTGPGGVVQGSKYAADRNHYRYPYRRGPPRYNYQNSSEGEKNEGSSESAPEG 225
 QY 181 QAOQRRPYRRRFPYMYRRPYGRPOYSNPPVQGEVMEGADNQGAGGQGRPVQNNMYRG 240
 Db 226 QAOQRRPYRRRFPYMYRRPYGRPOYSNPPVQGEVMEGADNQGAGGQGRPVQNNMYRG 285
 QY 241 YRPRFRGPPRQRPREDGNEEDKENQDGTQGGQPPQRRYRNFNRYRRRPNPKPQDG 300
 Db 286 YRPRFRGPPRQRPREDGNEEDKENQDGTQGGQPPQRRYRNFNRYRRRPNPKPQDG 345
 QY 301 KETKAADPPAENSAPAEQGGAE 324
 Db 346 KETKAADPPAENSAPAEQGGAE 369

RESULT 2

US-08-474-379C-32
 ; Sequence 32, Application US/08474379C
 ; Patent No. 5977305
 ; GENERAL INFORMATION:
 ; APPLICANT: Wigler, Michael H.
 ; APPLICANT: Colicelli, John J.
 ; TITLE OF INVENTION: CLONING BY COMPLEMENTATION AND RELATED
 ; TITLE OF INVENTION: PROCESSES
 ; NUMBER OF SEQUENCES: 88
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 233 South Wacker Drive/6300 Sears Tower
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/474,379C
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/511,715
 ; FILING DATE: 20-APR-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/206,188
 ; FILING DATE: 01-MAR-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/688,352
 ; FILING DATE: 19-APR-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Clough, David W.
 ; REGISTRATION NUMBER: 36,107
 ; REFERENCE/DOCKET NUMBER: 27866/32771
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (312) 474-6300
 ; TELEFAX: (312) 474-0448
 ; INFORMATION FOR SEQ ID NO: 32:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 369 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear

; MOLECULE TYPE: protein
 US-08-474-379C-32
 Query Match 100.0%; Score 1747; DB 2; Length 369;
 Best Local Similarity 100.0%; Pred. No. 1.1e-139;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSSEAEETQPPAAPALSAADTKGTTGSGAGSGGPGGLTSAAPAGGDKKVIATKVL 60
 Db 46 MSSEAEETQPPAAPALSAADTKGTTGSGAGSGGPGGLTSAAPAGGDKKVIATKVL 105
 QY 61 GTVKFNVNRYGGINRNDTKEDVVFHQTAKONPKYLSVGDGTVEFDVVEGEKGA 120
 Db 106 GTVKFNVNRYGGINRNDTKEDVVFHQTAKONPKYLSVGDGTVEFDVVEGEKGA 165
 QY 121 EAAVTGPGGVVQGSKYAADRNHYRYPYRRGPPRYNYQNSSEGEKNEGSSESAPEG 180
 Db 166 EAAVTGPGGVVQGSKYAADRNHYRYPYRRGPPRYNYQNSSEGEKNEGSSESAPEG 225
 QY 181 QAOQRRPYRRRFPYMYRRPYGRPOYSNPPVQGEVMEGADNQGAGGQGRPVQNNMYRG 240
 Db 226 QAOQRRPYRRRFPYMYRRPYGRPOYSNPPVQGEVMEGADNQGAGGQGRPVQNNMYRG 285
 QY 241 YRPRFRGPPRQRPREDGNEEDKENQDGTQGGQPPQRRYRNFNRYRRRPNPKPQDG 300
 Db 286 YRPRFRGPPRQRPREDGNEEDKENQDGTQGGQPPQRRYRNFNRYRRRPNPKPQDG 345
 QY 301 KETKAADPPAENSAPAEQGGAE 324
 Db 346 KETKAADPPAENSAPAEQGGAE 369

RESULT 3

US-09-146-249A-32
 ; Sequence 32, Application US/09146249A
 ; Patent No. 6069240
 ; GENERAL INFORMATION:
 ; APPLICANT: Wigler, Michael H.
 ; APPLICANT: Colicelli, John J.
 ; TITLE OF INVENTION: Cloning by Complementation and Related
 ; TITLE OF INVENTION: Processes
 ; NUMBER OF SEQUENCES: 85
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/146,249A
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/511,715
 ; FILING DATE: 20-APR-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Clough, David W.
 ; REGISTRATION NUMBER: 36,107
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312/474-6300
 ; TELEFAX: 312-474-0448
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 32:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 369 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear


```

; MOLECULE TYPE: protein
US-09-146-249A-32

Query Match      100.0%; Score 1747; DB 3; Length 369;
Best Local Similarity 100.0%; Pred. No. 1.1e-139;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSEAEETQPPAAPPALSAADTKPGTTGSGAGSGGPGGLTSAAPAGDKKVIATKVL 60
DB 46 MSSEAEETQPPAAPPALSAADTKPGTTGSGAGSGGPGGLTSAAPAGDKKVIATKVL 105
QY 61 GTVKWFNVRNGYGFINRNDTKEDVFVHQTAKNNPRKYLRSVGDGETVEFDVVEGEKGA 120
DB 106 GTVKWFNVRNGYGFINRNDTKEDVFVHQTAKNNPRKYLRSVGDGETVEFDVVEGEKGA 165
QY 121 EAAVTGPGGVQGSKYAADRNHYRRYPRRRGPPRPNYQNYQNSSEKNEGSAPAG 180
DB 166 EAAVTGPGGVQGSKYAADRNHYRRYPRRRGPPRPNYQNYQNSSEKNEGSAPAG 225
QY 181 QAOQRYPYRRRPPYMYRRPYGRPOYSNPPVQGEVMEGADNQGAGEQGRPVQRNMYRG 240
DB 226 QAOQRYPYRRRPPYMYRRPYGRPOYSNPPVQGEVMEGADNQGAGEQGRPVQRNMYRG 285
QY 241 YRPRFRGPPRQRPREDGNEEDKENOGDETQGGQPPORRYRNFYRRRRPENPKPDG 300
DB 286 YRPRFRGPPRQRPREDGNEEDKENOGDETQGGQPPORRYRNFYRRRRPENPKPDG 345
QY 301 KETKAADPPAENSAPAEQGGAE 324
DB 346 KETKAADPPAENSAPAEQGGAE 369

```

```

RESULT 4
US-08-206-188B-32
; Sequence 32, Application US/08206188B
; Patent No. 6100025
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: Cloning by Complementation and Related
; TITLE OF INVENTION: Processes
; NUMBER OF SEQUENCES: 84
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/206,188B
; FILING DATE: 01-MAR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,715
; FILING DATE: 20-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 369 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```

```

RESULT 5
PCT-US91-02714-31
; Sequence 31, Application PC/TUS9102714
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: Cloning by Complementation and Related
; TITLE OF INVENTION: Processes
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/02714
; FILING DATE: 19910419
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,715
; FILING DATE: 20-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25447
; REFERENCE/DOCKET NUMBER: 27805/30197
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 369 amino acids

```

```

;
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US91-02714-31

Query Match 100.0%; Score 1747; DB 5; Length 369;
Best Local Similarity 100.0%; Pred. No. 1.1e-139;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSAETQPPAAPALSAADTKGCTGSGAGSGPGGLTSAAPAGDGKKVIATKVL 60
Db 46 MSSAETQPPAAPALSAADTKGCTGSGAGSGPGGLTSAAPAGDGKKVIATKVL 105
QY 61 GTVKWFNVRNGYGFNRNDTKEDVFVHQTAKKNNPKYLSVGDGETVFDFVVEGEKGA 120
Db 106 GTVKWFNVRNGYGFNRNDTKEDVFVHQTAKKNNPKYLSVGDGETVFDFVVEGEKGA 165
QY 121 EAAVNTGGGVVQGSKYAADRNHYRRYPRRRGPPRYQNYQNSSEGEKNGESGAPEG 180
Db 166 EAAVNTGGGVVQGSKYAADRNHYRRYPRRRGPPRYQNYQNSSEGEKNGESGAPEG 225
QY 181 QAQRRPYRRRRFPYMRPGRPOYSNPPVOGEVMEGADNQGAGEQGRPVRRQNNYRG 240
Db 226 QAQRRPYRRRRFPYMRPGRPOYSNPPVOGEVMEGADNQGAGEQGRPVRRQNNYRG 285
QY 241 YRPRFRGPPRQRPREDGNEEDKENOGDETQGGQPPQRRYRNFNRRRPNPKPDG 300
Db 286 YRPRFRGPPRQRPREDGNEEDKENOGDETQGGQPPQRRYRNFNRRRPNPKPDG 345
QY 301 KETKAADPPAENSSAPAEQGA 324
Db 346 KETKAADPPAENSSAPAEQGA 369

```

```

RESULT 6
US-08-203-806B-13
; Sequence 13, Application US/08203806B
; Patent No. 5714575
; GENERAL INFORMATION:
; APPLICANT: Inouye, Masayori
; APPLICANT: Jones, Pamela
; APPLICANT: Etchegaray, Jean-Pierre
; APPLICANT: Weinig, Jian
; APPLICANT: Pollitt, N. Stephen
; APPLICANT: Goldstein, Joel
; TITLE OF INVENTION: Nucleic Acid Sequences, Stress-induced
; TITLE OF INVENTION: Proteins and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weiser & Associates
; STREET: 230 South Fifteenth Street, Suite 500
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/203,806B
; FILING DATE: 01-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiser, Gerard J.
; REGISTRATION NUMBER: 19,763
; REFERENCE/DOCKET NUMBER: 377,5998P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-875-8383
; TELEFAX: 215-875-8394
; TELEX: 834809 WEISTAK
; INFORMATION FOR SEQ ID NO: 13:

```

```

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 74 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-08-203-806B-13

Query Match 22.0%; Score 384; DB 1; Length 74;
Best Local Similarity 98.6%; Pred. No. 1.6e-25;
Matches 73; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 55 IATKVLGVTKWFNVRNGYGFNRNDTKEDVFVHQTAKKNNPKYLSVGDGETVFDFV 114
Db 1 IATKVLGVTKWFNVRNGYGFNRNDTKEDVFVHQTAKKNNPKYLSVGDGETVFDFV 60
QY 115 EGEKGAEEAANVTGP 128
Db 61 EGEKGAEEAANVTGP 74

RESULT 7
US-09-017-754A-13
; Sequence 13, Application US/09017754A
; Patent No. 6333191
; GENERAL INFORMATION:
; APPLICANT: Inouye, Masayori
; APPLICANT: Jones, Pamela
; APPLICANT: Etchegaray, Jean-Pierre
; APPLICANT: Weinig, Jian
; APPLICANT: Pollitt, N. Stephen
; APPLICANT: Goldstein, Joel
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES, STRESS-INDUCED
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCHNADER, HARRISON, SEGAL & LEWIS, LLP
; STREET: 1600 Market Street, Suite 3600
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/017,754A
; FILING DATE: 03-Feb-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiser, Gerard J.
; REGISTRATION NUMBER: 19,763
; REFERENCE/DOCKET NUMBER: 1113-cip2-div-00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-751-2427
; TELEFAX: 215-751-2658
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 74 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-017-754A-13

Query Match 22.0%; Score 384; DB 4; Length 74;
Best Local Similarity 98.6%; Pred. No. 1.6e-25;
Matches 73; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 55 IATKVLGVTKWFNVRNGYGFNRNDTKEDVFVHQTAKKNNPKYLSVGDGETVFDFV 114

```

```

Db 1 IATKVLGTWKFNVRNGYGFINRNDTKEDVHVHQTAKKNNPKYLSVGDGETVEFDV 60
QY 115 EGEKGAEAAVNTGP 128
Db 61 EGEKGAEAAVNTGP 74

RESULT 8
US-09-489-039A-13476
; Sequence 13476, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709 2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13476
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13476

Query Match 9.8%; Score 172; DB 4; Length 131;
Best Local Similarity 43.2%; Pred. No. 2.5e-07;
Matches 32; Conservative 15; Mismatches 23; Indels 4; Gaps 1;

QY 52 KVIATKVLGTWKFNVRNGYGFINRNDTKEDVHVHQTAKKNNPKYLSVGDGETVEF 111
Db 59 KVIATKVLGTWKFNVRNGYGFINRNDTKEDVHVHQTAKKNNPKYLSVGDGETVEF 114
QY 112 DVVEGEGKAEAAVNTGP 125
Db 115 EITNGAKGPSAAVNTGP 128

RESULT 9
US-09-134-001C-5476
; Sequence 5476, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5476
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5476

Query Match 9.6%; Score 168.5; DB 4; Length 82;
Best Local Similarity 52.3%; Pred. No. 2.8e-07;
Matches 34; Conservative 13; Mismatches 13; Indels 5; Gaps 2;

QY 61 GTVKWFNVRNGYGFINRNDTKEDVHVHQTAKKNNPKYLSVGDGETVEFDVVEGEGA 120
Db 20 GTVKWFNVRNGYGFINRNDTKEDVHVHQTAKKNNPKYLSVGDGETVEFDVVEGEGA 74
QY 121 EAAVNTGP 125

```

```

Db 75 QAAVNTGP 79

RESULT 10
US-09-252-991A-19664
; Sequence 19664, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19664
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19664

Query Match 9.4%; Score 164.5; DB 4; Length 116;
Best Local Similarity 47.9%; Pred. No. 9.4e-07;
Matches 35; Conservative 14; Mismatches 19; Indels 5; Gaps 2;

QY 54 VIATKVLGTWKFNVRNGYGFINRNDTKEDVHVHQTAKKNNPKYLSVGDGETVEFDV 113
Db 47 VMADEVECTVKFNVRNGYGFINRNDTKEDVHVHQTAKKNNPKYLSVGDGETVEFDV 101
QY 114 VEGEGKAEAAVNTGP 126
Db 102 IQGQKGLQAEVNTGP 114

RESULT 11
US-09-107-532A-5955
; Sequence 5955, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Denise
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007

```

```

; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5955:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...86
; SEQUENCE DESCRIPTION: SEQ ID NO: 5955:
US-09-107-532A-5955

Query Match          9.4%; Score 163.5; DB 4; Length 86;
Best Local Similarity 40.5%; Pred. No. 7.8e-07;
Matches 30; Conservative 22; Mismatches 17; Indels 5; Gaps 2;

QY 52 KVLATKVLGVNVRNGYGFIRNNDTKEDVVFHQTAKKNPKYLSVGDGETVFEF 111
Db 4 RKVSALTGGLVWFDNKGKGFISYDDT-EEIFVHFTALEEG----FKTLEENQVVEF 58

QY 112 DWVEGEGAAAVY 125
Db 59 EIIENRGTOAAHV 72

RESULT 12
US-08-203-806B-4
; Sequence 4, Application US/08203806B
; Patent No. 5714575
; GENERAL INFORMATION:
; APPLICANT: Inouye, Masayori
; APPLICANT: Jones, Pamela
; APPLICANT: Etchegaray, Jean-Pierre
; APPLICANT: Weinig, Jian
; APPLICANT: Pollitt, N. Stephen
; APPLICANT: Goldstein, Joel
; TITLE OF INVENTION: Nucleic Acid Sequences, Stress-induced
; TITLE OF INVENTION: Proteins and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weiser & Associates
; STREET: 230 South Fifteenth Street, Suite 500
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/203,806B
; FILING DATE: 01-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiser, Gerard J.
; REGISTRATION NUMBER: 19,763
; REFERENCE/DOCKET NUMBER: 377,5998P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-875-8383
; TELEFAX: 215-875-8394
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORGANISM: Escherichia coli
; ORGANISM: Escherichia coli
US-08-203-806B-9

Query Match          9.3%; Score 163; DB 1; Length 69;
Best Local Similarity 44.9%; Pred. No. 6.6e-07;
Matches 31; Conservative 14; Mismatches 20; Indels 4; Gaps 1;

QY 58 KVLGVNVRNGYGFIRNNDTKEDVVFHQTAKKNPKYLSVGDGETVFEFVVVEGE 117
Db 3 KIKGVNVRNGYGFIRNNDTKEDVVFHQTAKKNPKYLSVGDGETVFEFVVVEGE 58

QY 118 KGAEAAVNT 126
Db 59 KGPAAVNT 67

RESULT 13
US-08-203-806B-9
; Sequence 9, Application US/08203806B
; Patent No. 5714575
; GENERAL INFORMATION:
; APPLICANT: Inouye, Masayori
; APPLICANT: Jones, Pamela
; APPLICANT: Etchegaray, Jean-Pierre
; APPLICANT: Weinig, Jian
; APPLICANT: Pollitt, N. Stephen
; APPLICANT: Goldstein, Joel
; TITLE OF INVENTION: Nucleic Acid Sequences, Stress-induced
; TITLE OF INVENTION: Proteins and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weiser & Associates
; STREET: 230 South Fifteenth Street, Suite 500
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/203,806B
; FILING DATE: 01-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiser, Gerard J.
; REGISTRATION NUMBER: 19,763
; REFERENCE/DOCKET NUMBER: 377,5998P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-875-8383
; TELEFAX: 215-875-8394
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; ORGANISM: Escherichia coli
US-08-203-806B-9

Query Match          9.3%; Score 163; DB 1; Length 69;
Best Local Similarity 44.9%; Pred. No. 6.6e-07;
Matches 31; Conservative 14; Mismatches 20; Indels 4; Gaps 1;

QY 58 KVLGVNVRNGYGFIRNNDTKEDVVFHQTAKKNPKYLSVGDGETVFEFVVVEGE 117
Db 3 KIKGVNVRNGYGFIRNNDTKEDVVFHQTAKKNPKYLSVGDGETVFEFVVVEGE 58

QY 118 KGAEAAVNT 126
Db 59 KGPAAVNT 67
```

```

Db      59 KGPAANVT 67

RESULT 14
US-09-017-754A-4
; Sequence 4, Application US/09017754A
; Patent No. 6333191
; GENERAL INFORMATION:
; APPLICANT: INOUE, MASAYORI
; JONES, PAMELA
; ETCHEGARAY, JEAN-PIERRE
; WEINING, JIANG
; POLITT, N. STEPHEN
; GOLDSTEIN, JOEL
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES, STRESS-INDUCED
; PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: SCHNADER, HARRISON, SEGAL & LEWIS, LLP
; STREET: 1600 Market Street, Suite 3600
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/017,754A
; FILING DATE: 03-Feb-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiser, Gerard J.
; REGISTRATION NUMBER: 19,763
; REFERENCE/DOCKET NUMBER: 1113-cip2-div-00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-751-2427
; TELEFAX: 215-751-2658
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-017-754A-4

Query Match      9.3%; Score 163; DB 4; Length 69;
Best Local Similarity 44.9%; Pred. No. 6.6e-07;
Matches 31; Conservative 14; Mismatches 20; Indels 4; Gaps 1;

Qy      58 KVLGTVKWVNRNGYGFINRNDTKEDVHQTAKKNPKYLSVGGDGETVEFDVVEGE 117
      |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3 KIKGVKWFNKGFGFITPADGSKDVVHFSAIQNG-----FKTLAEGQNVFEIQDQG 58
      |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy      118 KGAEAAVNT 126
      |||||
Db      59 KGPAANVT 67

Search completed: August 25, 2004, 00:27:41
Job time : 34 secs

```

```

Db      59 KGPAANVT 67

RESULT 15
US-09-017-754A-9
; Sequence 9, Application US/09017754A
; Patent No. 6333191
; GENERAL INFORMATION:
; APPLICANT: INOUE, MASAYORI
; JONES, PAMELA
; ETCHEGARAY, JEAN-PIERRE
; WEINING, JIANG
; POLITT, N. STEPHEN
; GOLDSTEIN, JOEL
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES, STRESS-INDUCED
; PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: SCHNADER, HARRISON, SEGAL & LEWIS, LLP
; STREET: 1600 Market Street, Suite 3600
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/017,754A
; FILING DATE: 03-Feb-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiser, Gerard J.
; REGISTRATION NUMBER: 19,763
; REFERENCE/DOCKET NUMBER: 1113-cip2-div-00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-751-2427
; TELEFAX: 215-751-2658
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-017-754A-9

Query Match      9.3%; Score 163; DB 4; Length 69;
Best Local Similarity 44.9%; Pred. No. 6.6e-07;
Matches 31; Conservative 14; Mismatches 20; Indels 4; Gaps 1;

Qy      58 KVLGTVKWVNRNGYGFINRNDTKEDVHQTAKKNPKYLSVGGDGETVEFDVVEGE 117
      |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      3 KIKGVKWFNKGFGFITPADGSKDVVHFSAIQNG-----FKTLAEGQNVFEIQDQG 58
      |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy      118 KGAEAAVNT 126
      |||||
Db      59 KGPAANVT 67

Search completed: August 25, 2004, 00:27:41
Job time : 34 secs

```

This Page Blank (uspio)

Sequence 142, App
Sequence 142, App
Sequence 221, App
Sequence 4157, App
Sequence 221, App
Sequence 39, App
Sequence 143, App
Sequence 4161, App
Sequence 143, App
Sequence 62338, A
Sequence 40158, A
Sequence 40165, A
Sequence 42358, A
Sequence 41792, A
Sequence 42359, A
Sequence 41771, A
Sequence 256768, A
Sequence 41411, A
Sequence 39822, A
Sequence 256772, A
Sequence 42357, A
Sequence 256774, A
Sequence 41777, A
Sequence 39486, A
Sequence 245, App
Sequence 256, App
Sequence 155394, A
Sequence 1041, App
Sequence 110, App

16 495 28.3 303 9 US-09-764-846-142
17 495 28.3 303 14 US-10-091-483-142
18 476.5 27.3 253 9 US-09-764-846-221
19 476.5 27.3 253 10 US-09-764-891-4157
20 476.5 27.3 253 14 US-10-091-483-221
21 350 20.0 66 13 US-10-028-415-39
22 288 16.5 167 9 US-09-764-846-143
23 288 16.5 167 10 US-09-764-891-4161
24 288 16.5 167 14 US-10-091-483-143
25 199.5 11.4 174 12 US-10-425-114-62338
26 191.5 11.0 154 12 US-10-425-114-40158
27 191.5 11.0 154 12 US-10-425-114-40165
28 191.5 11.0 206 12 US-10-425-114-42358
29 191.5 11.0 206 12 US-10-425-114-41792
30 191.5 11.0 208 12 US-10-425-114-42359
31 191.5 11.0 210 12 US-10-425-114-41779
32 191.5 11.0 229 12 US-10-425-114-41771
33 189.5 10.8 148 12 US-10-424-599-256768
34 189.5 10.8 156 12 US-10-425-114-41411
35 189.5 10.8 202 12 US-10-424-599-256772
36 189.5 10.8 208 12 US-10-425-114-42357
37 189.5 10.8 201 12 US-10-424-599-256774
38 187.5 10.7 206 12 US-10-425-114-41777
39 187.5 10.7 328 12 US-10-425-114-39486
40 187.5 10.5 151 12 US-10-220-120-245
41 183 10.5 151 12 US-10-363-829-256
42 183 10.5 151 16 US-10-437-963-155394
43 181.5 10.4 241 16 US-10-374-780A-1041
44 178.5 10.2 197 15 US-10-225-066A-110
45 178.5 10.2 201 12 US-10-225-066A-110

ALIGNMENTS

RESULT 1
US-10-028-415-40
; Sequence 40, Application US/10028415
; Publication No. US20020151063A1
; GENERAL INFORMATION:
; APPLICANT: Lasham, Amette
; APPLICANT: Watson, James D.
; TITLE OF INVENTION: Methods for Modulating Apoptotic Cell
; TITLE OF INVENTION: Death
; FILE REFERENCE: 11000-1004c3
; CURRENT APPLICATION NUMBER: US/10/028,415
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: PCT/NZ01/00286
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: US 09/724,809
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/036,004
; PRIOR FILING DATE: 1998-03-04
; PRIOR APPLICATION NUMBER: US 08/713,557
; PRIOR FILING DATE: 1996-08-30
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Human
US-10-028-415-40

Query Match 100.0%; Score 1747; DB 13; Length 324;
Best Local Similarity 100.0%; Pred. No. 9.5e-127;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSSEATQPPAAPPAALSAADTKPGTTGSGAGSGGGGLTSAAPAGGDKKVIATKVL 60
Db 1 MSSEATQPPAAPPAALSAADTKPGTTGSGAGSGGGGLTSAAPAGGDKKVIATKVL 60
QY 61 GTVKFNVRNGVFINRNDTKEDVFVHQTAKNNPRKVLRSVGDETFVDFVDSGEKA 120

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 00:26:26 ; Search time 127 Seconds
(without alignments)
801.722 Million cell updates/sec

Title: US-10-028-415-40

Perfect score: 1747

Sequence: 1 MSSEATQPPAAPPAAL.....AADPPAENSSAPEAEQGGAE 324

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1295152 seqs, 31425058 residues

Total number of hits satisfying chosen parameters: 1295152

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PTCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PTCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1747	100.0	324	13	US-10-028-415-40 Sequence 40, Appl
2	1707	97.7	322	12	US-10-388-215-57 Sequence 57, Appl
3	1701	97.4	322	12	US-10-388-215-56 Sequence 56, Appl
4	1699	97.3	322	12	US-10-388-215-2 Sequence 2, Appl
5	1699	97.3	322	12	US-10-388-215-51 Sequence 51, Appl
6	1699	97.3	322	12	US-10-388-215-53 Sequence 53, Appl
7	1690	96.7	322	12	US-10-388-215-55 Sequence 55, Appl
8	1688	96.6	322	12	US-10-388-215-58 Sequence 58, Appl
9	1687	96.6	317	9	US-09-972-331-1 Sequence 1, Appl
10	1687	96.6	317	12	US-10-388-215-4 Sequence 4, Appl
11	1687	96.6	317	14	US-10-023-969-3 Sequence 3, Appl
12	773	44.2	412	15	US-10-440-464-124 Sequence 124, Appl
13	750.5	43.0	303	15	US-10-408-765A-1896 Sequence 1896, Ap
14	731	41.8	272	9	US-09-925-300-1697 Sequence 1697, Ap
15	640	36.6	164	12	US-10-424-599-280985 Sequence 280985,

```

Db      61 GTVKWFNVRNGYGFINRNDTKEDVFVHQTAKKNNPKYLSVGDGETVEFDVVEGEKA 120
QY      121 EAANVTGGVPGVQGSKYAADRNHYRRYPRRRGPPRYNQYQNSSEKNEGSESAPEG 180
Db      121 EAANVTGGVPGVQGSKYAADRNHYRRYPRRRGPPRYNQYQNSSEKNEGSESAPEG 180
QY      181 QAQRRPYRRRRFPYMYRRPYGRRPOYSNPPVQGEVMEGADNOGAGEQGRPVQNMVYRG 240
Db      181 QAQRRPYRRRRFPYMYRRPYGRRPOYSNPPVQGEVMEGADNOGAGEQGRPVQNMVYRG 240
QY      241 YRPRFRGPPRQRPREDGNEEDKENOGDETQGGQPPQRRYRNFNRYRRRPNPKPDG 300
Db      241 YRPRFRGPPRQRPREDGNEEDKENOGDETQGGQPPQRRYRNFNRYRRRPNPKPDG 300
QY      301 KETKAADPPAENSSAPEAQOGAE 324
Db      301 KETKAADPPAENSSAPEAQOGAE 324

RESULT 2
US-10-388-215-57
; Sequence 57, Application US/10388215
; Publication No. US20030223975A1
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas K.
; APPLICANT: Fukada, Toshiyuki
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATION OF PTP-1B
; FILE REFERENCE: 200125.440
; CURRENT APPLICATION NUMBER: US/10388,215
; CURRENT FILING DATE: 2003-03-11
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-388-215-57

```

```

Query Match      97.7%; Score 1707; DB 12; Length 322;
Best Local Similarity 98.5%; Pred. No. 1.2e-123;
Matches 319; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

```

```

QY      1 MSSEAEETQPPAAP--AAALSADTKPGSTGAGSGGPGGLTSAAPAGDKKVIATKVL 60
Db      1 MSSEAEETQPPAAP--AAALSADTKPGSTGAGSGGPGGLTSAAPAGDKKVIATKVL 58
QY      61 GTVKWFNVRNGYGFINRNDTKEDVFVHQTAKKNNPKYLSVGDGETVEFDVVEGEKA 120
Db      59 GTVKWFNVRNGYGFINRNDTKEDVFVHQTAKKNNPKYLSVGDGETVEFDVVEGEKA 118
QY      121 EAANVTGGVPGVQGSKYAADRNHYRRYPRRRGPPRYNQYQNSSEKNEGSESAPEG 180
Db      119 EAANVTGGVPGVQGSKYAADRNHYRRYPRRRGPPRYNQYQNSSEKNEGSESAPEG 178
QY      181 QAQRRPYRRRRFPYMYRRPYGRRPOYSNPPVQGEVMEGADNOGAGEQGRPVQNMVYRG 240
Db      179 QAQRRPYRRRRFPYMYRRPYGRRPOYSNPPVQGEVMEGADNOGAGEQGRPVQNMVYRG 238
QY      241 YRPRFRGPPRQRPREDGNEEDKENOGDETQGGQPPQRRYRNFNRYRRRPNPKPDG 300
Db      239 YRPRFRGPPRQRPREDGNEEDKENOGDETQGGQPPQRRYRNFNRYRRRPNPKPDG 298
QY      301 KETKAADPPAENSSAPEAQOGAE 324
Db      299 KETKAADPPAENSSAPEAQOGAE 322

```

```

RESULT 3
US-10-388-215-56
; Sequence 56, Application US/10388215
; Publication No. US20030223975A1
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas K.

```

```

; APPLICANT: Fukada, Toshiyuki
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATION OF PTP-1B
; FILE REFERENCE: 200125.440
; CURRENT APPLICATION NUMBER: US/10388,215
; CURRENT FILING DATE: 2003-03-11
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-388-215-56

```

```

Query Match      97.4%; Score 1701; DB 12; Length 322;
Best Local Similarity 98.1%; Pred. No. 3.3e-123;
Matches 318; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

```

```

QY      1 MSSEAEETQPPAAP--AAALSADTKPGSTGAGSGGPGGLTSAAPAGDKKVIATKVL 60
Db      1 MSSEAEETQPPAAP--AAALSADTKPGSTGAGSGGPGGLTSAAPAGDKKVIATKVL 58
QY      61 GTVKWFNVRNGYGFINRNDTKEDVFVHQTAKKNNPKYLSVGDGETVEFDVVEGEKA 120
Db      59 GTVKWFNVRNGYGFINRNDTKEDVFVHQTAKKNNPKYLSVGDGETVEFDVVEGEKA 118
QY      121 EAANVTGGVPGVQGSKYAADRNHYRRYPRRRGPPRYNQYQNSSEKNEGSESAPEG 180
Db      119 EAANVTGGVPGVQGSKYAADRNHYRRYPRRRGPPRYNQYQNSSEKNEGSESAPEG 178
QY      181 QAQRRPYRRRRFPYMYRRPYGRRPOYSNPPVQGEVMEGADNOGAGEQGRPVQNMVYRG 240
Db      179 QAQRRPYRRRRFPYMYRRPYGRRPOYSNPPVQGEVMEGADNOGAGEQGRPVQNMVYRG 238
QY      241 YRPRFRGPPRQRPREDGNEEDKENOGDETQGGQPPQRRYRNFNRYRRRPNPKPDG 300
Db      239 YRPRFRGPPRQRPREDGNEEDKENOGDETQGGQPPQRRYRNFNRYRRRPNPKPDG 298
QY      301 KETKAADPPAENSSAPEAQOGAE 324
Db      299 KETKAADPPAENSSAPEAQOGAE 322

```

```

RESULT 4
US-10-388-215-2
; Sequence 2, Application US/10388215
; Publication No. US20030223975A1
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas K.
; APPLICANT: Fukada, Toshiyuki
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATION OF PTP-1B
; FILE REFERENCE: 200125.440
; CURRENT APPLICATION NUMBER: US/10388,215
; CURRENT FILING DATE: 2003-03-11
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-388-215-2

```

```

Query Match      97.3%; Score 1699; DB 12; Length 322;
Best Local Similarity 98.1%; Pred. No. 4.8e-123;
Matches 318; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

```



```

121 QY EAANVTGGVPGVGSKYAADRNHYRYPRRRGPPPNYQONTQNSGSEKNEGSESAPG 180
Db EAANVTGGVPGVGSKYAADRNHYRYPRRRGPPPNYQONTQNSGSEKNEGSESAPG 178
QY QAQORRPPYRRRRFPYYMRRPYPQRRPOYSNPPVQGWMEGADNOGAGQGRRPVRQNMVYG 240
Db QAQORRPPYRRRRFPYYMRRPYPQRRPOYSNPPVQGWMEGADNOGAGQGRRPVRQNMVYG 238
QY YRPRFRGPPRQOPFRDGNHEEKENOGDETQCGQPQRRYRNRYRRRRPKNPKPDG 300
Db YRPRFRGPPRQOPFRDGNHEEKENOGDETQCGQPQRRYRNRYRRRRPKNPKPDG 298
QY KETKAADPPPAENSSAPAEQGGAE 324
Db KETKAADSPAENSSAPAEQGGAE 322

RESULT 5
US-10-388-215-51
; Sequence 51, Application US/10388215
; Publication No. US20030223975A1
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas K.
; APPLICANT: Fukada, Toshiyuki
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATION OF PTP-1B
; FILE REFERENCE: 200125.440
; CURRENT APPLICATION NUMBER: US/10/388,215
; CURRENT FILING DATE: 2003-03-11
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-388-215-51

```

Query Match	97.3%;	Score 1699;	DB 12;	Length 322;
Best Local Similarity	98.1%;	Pred. No. 4.8e-123;		
Matches 318;	Conservative	1;	Mismatches 3;	Indels 2;
Gaps 1;				

1	MSSAEATQPPAAPAPALSAADTPGTGSGAGSGGGLTSAAPAGGDKKIATKVL	60	Qy
1	MSSAEATQPPAAPAPALSAADTPGTGSGAGSGGGLTSAAPAGGDKKIATKVL	58	Db
61	GTVKWVNRNGYGFINRNDTKEDVVFHQIAIKNNPRKYLRSVGDETVEFDVVEGEKGA	120	Qy
59	GTVKWVNRNGYGFINRNDTKEDVVFHQIAIKNNPRKYLRSVGDETVEFDVVEGEKGA	118	Db
121	EAAVNTGGGVVQGSKYAADRHYYRRYPRRRGPPRNYQNTYQNSSESGKNEGSESAPG	180	Qy
119	EAAVNTGGGVVQGSKYAADRHYYRRYPRRRGPPRNYQNTYQNSSESGKNEGSESAPG	178	Db
181	QAQQRYPYRRRRPPPYMYR3PYGRRPOYSNPVQGEVMEGADNQAGSGGQRPVRCNMYRG	240	Qy
179	QAQQRYPYRRRRPPPYMYR3PYARRPOYSNPVQGEVMEGADNQAGSGGQRPVRCNMYRG	238	Db
241	YRPRFRGPPROROPREDGNEEBKQNGDETCQQQPORRYRNNYRRRRPENPKPODG	300	Qy
239	YRPRFRGPPROROPREDGNEEBKQNGDETCQQQPORRYRNNYRRRRPENPKPODG	298	Db
301	KETKAADPPPAENSSAPAEQGGAE	324	Qy
299	KETKAADSPAENSSAPAEQGGAE	322	Db

```

RESULT 6
US 10 388-215-53
Sequence 53 Application US/10388215
Publication No. US2000223975A1
GENERAL INFORMATION:
APPLICANT: Tonks, Nicholas K.
INVENTOR: Fukada, Toshiyuki
TITLE OF INVENTION: TRANSCRIPTIONAL REGULATION OF PT2-1R

```

Db 119 EAAVTGPGVPGVQSGKYAADRNHYRRYRRRGGPPPNYQNYQNSGEGKNGESAPG 178
 QY 181 QAOQRRPYRRRPPYRRPYRRPQYNSNPPVQGEVMEGADNQGAGGQGRPVQNNMYRG 240
 Db 179 QAOQRRPYRRRPPYRRPYRRPQYNSNPPVQGEVMEGADNQGAGGQGRPVQNNMYRG 238
 QY 241 YPRFRGPPRQRPREDGNEEDKNGDETCGQPPORRYRRNFYRRRPPENPKPDG 300
 Db 239 YPRFRGPPRQRPREDGNEEDKNGDETCGQPPORRYRRNFYRRRPPENPKPDG 298
 QY 301 KETKAADPPAENSSAPEAQGGAE 324
 Db 299 KETKAADPPAENSSAPEAQGGAE 322

RESULT 8

US-10-388-215-58
 ; Sequence 58, Application US/10388215
 ; Publication No. US20030223975A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tonks, Nicholas K.
 ; APPLICANT: Fukada, Toshiyuki
 ; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATION OF PTP-1B
 ; FILE REFERENCE: 200125.440
 ; CURRENT APPLICATION NUMBER: US/10/388,215
 ; CURRENT FILING DATE: 2003-03-11
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 58
 ; LENGTH: 322
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 US-10-388-215-58

Query Match 96.6%; Score 1688; DB 12; Length 322;
 Best Local Similarity 97.5%; Pred. No. 3.4e-122;
 Matches 316; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 1 MSSEAEATQPPAAPALSAADTKPGTTGSGAGSGGPGGLTSAAPAGGDKKIATKVL 60
 Db 1 MSSEAEATQPPAAP--AAALSAADTKPGT:GSGAGSGGPGGLTSAAPAGGDKKIATKVL 58
 QY 61 GTVKWFNVNRNGYGFINRNDTKEDVFVHQTAKKNNPKYLRVSGDGETVEFDVVEGEKGA 120
 Db 59 GTVKWFNVNRNGYGFINRNDTKEDVFVHQTAKKNNPKYLRVSGDGETVEFDVVEGEKGA 118
 QY 121 EAAVTGPGVPGVQSGKYAADRNHYRRYRRRGGPPPNYQNYQNSGEGKNGESAPG 180
 Db 119 EAAVTGPGVPGVQSGKYAADRNHYRRYRRRGGPPPNYQNYQNSGEGKNGESAPG 178
 QY 181 QAOQRRPYRRRPPYRRPYRRPQYNSNPPVQGEVMEGADNQGAGGQGRPVQNNMYRG 240
 Db 179 QAOQRRPYRRRPPYRRPYRRPQYNSNPPVQGEVMEGADNQGAGGQGRPVQNNMYRG 238
 QY 241 YPRFRGPPRQRPREDGNEEDKNGDETCGQPPORRYRRNFYRRRPPENPKPDG 300
 Db 239 YPRFRGPPRQRPREDGNEEDKNGDETCGQPPORRYRRNFYRRRPPENPKPDG 298
 QY 301 KETKAADPPAENSSAPEAQGGAE 324
 Db 299 KETKAADPPAENSSAPEAQGGAE 322

RESULT 9

US-09-972-331-1
 ; Sequence 1, Application US/09972331
 ; Patent No. US20020091083A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HIGASHI, KIYOSHI
 ; APPLICANT: KOMATSU, KENGO
 ; TITLE OF INVENTION: DNA-BINDING PROTEIN YB-1-CONTAINING
 ; TITLE OF INVENTION: COLLAGEN ACCUMULATION INHIBITORS
 ; FILE REFERENCE: 7372/72170

; CURRENT APPLICATION NUMBER: US/09/972,331
 ; CURRENT FILING DATE: 2001-12-28
 ; PRIOR APPLICATION NUMBER: JP 2000/310624
 ; PRIOR FILING DATE: 2000-10-11
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 317
 ; TYPE: PRT
 ; ORGANISM: HOMO SAPIENS
 US-09-972-331-1

Query Match 96.6%; Score 1687; DB 9; Length 317;
 Best Local Similarity 99.7%; Pred. No. 4e-122;
 Matches 312; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSEAEATQPPAAPALSAADTKPGTTGSGAGSGGPGGLTSAAPAGGDKKIATKVL 60
 Db 1 MSSEAEATQPPAAPALSAADTKPGTTGSGAGSGGPGGLTSAAPAGGDKKIATKVL 60
 QY 61 GTVKWFNVNRNGYGFINRNDTKEDVFVHQTAKKNNPKYLRVSGDGETVEFDVVEGEKGA 120
 Db 61 GTVKWFNVNRNGYGFINRNDTKEDVFVHQTAKKNNPKYLRVSGDGETVEFDVVEGEKGA 120
 QY 121 EAAVTGPGVPGVQSGKYAADRNHYRRYRRRGGPPPNYQNYQNSGEGKNGESAPG 180
 Db 121 EAAVTGPGVPGVQSGKYAADRNHYRRYRRRGGPPPNYQNYQNSGEGKNGESAPG 180
 QY 181 QAOQRRPYRRRPPYRRPYRRPQYNSNPPVQGEVMEGADNQGAGGQGRPVQNNMYRG 240
 Db 181 QAOQRRPYRRRPPYRRPYRRPQYNSNPPVQGEVMEGADNQGAGGQGRPVQNNMYRG 240
 QY 241 YPRFRGPPRQRPREDGNEEDKNGDETCGQPPORRYRRNFYRRRPPENPKPDG 300
 Db 241 YPRFRGPPRQRPREDGNEEDKNGDETCGQPPORRYRRNFYRRRPPENPKPDG 300
 QY 301 KETKAADPPAENS 313
 Db 301 KETKAADPPAENS 313

RESULT 10

US-10-388-215-4
 ; Sequence 4, Application US/10388215
 ; Publication No. US20030223975A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tonks, Nicholas K.
 ; APPLICANT: Fukada, Toshiyuki
 ; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATION OF PTP-1B
 ; FILE REFERENCE: 200125.440
 ; CURRENT APPLICATION NUMBER: US/10/388,215
 ; CURRENT FILING DATE: 2003-03-11
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 317
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-388-215-4

Query Match 96.6%; Score 1687; DB 12; Length 317;
 Best Local Similarity 99.7%; Pred. No. 4e-122;
 Matches 312; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSEAEATQPPAAPALSAADTKPGTTGSGAGSGGPGGLTSAAPAGGDKKIATKVL 60
 Db 1 MSSEAEATQPPAAPALSAADTKPGTTGSGAGSGGPGGLTSAAPAGGDKKIATKVL 60
 QY 61 GTVKWFNVNRNGYGFINRNDTKEDVFVHQTAKKNNPKYLRVSGDGETVEFDVVEGEKGA 120
 Db 61 GTVKWFNVNRNGYGFINRNDTKEDVFVHQTAKKNNPKYLRVSGDGETVEFDVVEGEKGA 120
 QY 121 EAAVTGPGVPGVQSGKYAADRNHYRRYRRRGGPPPNYQNYQNSGEGKNGESAPG 180


```
; SOFTWARE: FastSeq for Windows-Version 4.0
; SEQ ID NO 1896
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1896

Query Match      43.0%; Score 750.5; DB 16; Length 303;
Best Local Similarity 49.9%; Pred. No. 6.9e-50;
Matches 181; Conservative 15; Mismatches 68; Indels 99; Gaps 14;

QY 1 MSSEAE-----TOQPPAAPAALSAADTKP-CITGSGAG-----SGGPGGLT 43
Db 1 MSEAGEATTTTTLPGAPTEAAAAAQQDPAPKSPVSGGAPAAAPAAHVAGNFGG-- 58

QY 44 SAAPAG-----GDKVIATKVLGTVKFNVRNGYGFNRNDTKEDVFV 86
Db 59 DAAPAAATGTAATAAAGSDEAEKKVLATKVLGTVKFNVRNGYGFNRNDTKEDVFV 118

QY 87 HTAIKKNPKYLSVGDGETVFVVEGEGAEAAANTGPGVPVQGSKYAADRNHYR 146
Db 119 HTAIKKNPKYLSIGDGETVFVVEGEGAEAAANTGPGVPVQGSKYAADRNHYR 178

QY 147 R--YPRRGPRNTYQNYQNSSEKNEGSSEAPGQAOQRRPYRRRPPYMYRRPYGR 204
Db 179 RGYGRRRRGPPRN-----AGEIGEMKDGVPPEG-AQLQGPVHR----- 214

QY 205 RPQYSNPPVQGVMEGADNQAGEGQGRPVQNMVRYRPRFR-RGPRQRPQPDGNEED 263
Db 215 -----NPT-----YRPRYRSGRPPRPAPAVGEAED 241

QY 264 KENQDETQGGQPPQRR-YRNFNRYRR-RPENPKPDGKHETKAADPPAENSAPAEQG 321
Db 242 KENQATSGNPQSVRGYRRPYNYRRPPRPPNAPSQDGKEAKAGEAPTEN-PAPPTQOS 300

QY 322 GAE 324
Db 301 SAE 303

RESULT 14
US-09-925-300-1697
; Sequence 1697, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Ruben,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P4101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1897
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (256)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (258)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (262)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (263)
```

```
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (267)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1697

Query Match      41.8%; Score 731; DB 9; Length 272;
Best Local Similarity 53.5%; Pred. No. 1.9e-48;
Matches 166; Conservative 15; Mismatches 59; Indels 70; Gaps 10;

QY 17 APALSA-----DTKPGTTGSGAGSGGPGGLTSAAPAGGDKKVIATKVLGTVKFNVRNG 71
Db 2 APAAHVAGNPGGDAAPAAATGTAAAA---SLATAAGSEDAEKKVLATKVLGTVKFNVRNG 58

QY 72 YGFNRNDTKEDVFVHTAIKKNPKYLSVGDGETVFVVEGEGAEAAANTGPGGV 131
Db 59 YGFNRNDTKEDVFVHTAIKKNPKYLSVGDGETVFVVEGEGAEAAANTGPDGV 118

QY 132 PVQGSKYAADRNHYR--YPRRGPRNTYQNYQNSSEKNEGSSEAPGQAOQRRPYR 189
Db 119 PVESRYAADRRYRERGYYGRRRGPPN-----AGEIGEMKDGVPPEG-AQLQGPVH 168

QY 190 RRRFPFYMYRRPYRGRPOYSNPPVQGVMEGADNQAGEGQGRPVQNMVRYRPRFR-RG 248
Db 169 R-----NPT-----YRPRYRSGR 181

QY 249 PPRQRPQPDGNEEDKENQDETQGGQPPQRR-YRNFNRYRR-RPENPKPDGKHETKAA 306
Db 182 PPRPRPAPAVGEADKENQATSGNPQSVRGYRRPYNYRRPPRPPNAPSQDGKEAKAG 241

QY 307 DPPAENSAP 316
Db 242 EAPTENPAPP 251

RESULT 15
US-10-424-599-280985
; Sequence 280985, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 280985
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_95751C.1.pep
US-10-424-599-280985

Query Match      36.6%; Score 640; DB 12; Length 164;
Best Local Similarity 99.2%; Pred. No. 1.1e-41;
Matches 123; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSEAEQCPAPPAAPALSAADTKGTTGSGAGSGGPGGLTSAAPAGGDKKVIATKVL 60
Db 41 MSSEAEQCPAPPAAPPAALSAADTKGTTGSGAGSGGPGGLTSAAPAGGDKKVIATKVL 100

QY 61 GTVKFNVRNGYGFNRNDTKEDVFVHTAIKKNPKYLSVGDGETVFVVEGEGKGA 120
Db 101 GTVKFNVRNGYGFNRNDTKEDVFVHTAIKKNPKYLSVGDGETVFVVEGEGKGA 160

QY 121 EAA 124
Db 161 EAA 164
```

Search completed: August 25, 2004, 00:38:03
Job time : 129 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 00:18:50 ; Search time 40 Seconds
(without alignments)
779.151 Million cell updates/sec

Title: US-10-028-415-40
Perfect score: 1747
Sequence: 1 MSSEATQPPAAPPAAL.....AADPPAENSAPEAQGAE 324

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1747	100.0	324	1 I39382	y box-binding prot
2	1747	100.0	324	1 J02922	y box-binding prot
3	1728	98.9	324	1 A55971	y box-binding prot
4	1701	97.4	322	1 A58195	y box-binding prot
5	1699	97.3	322	1 A23677	y box-binding prot
6	1589.5	91.0	321	1 A48136	nuclease sensitive
7	1383	79.2	321	1 S34426	y box-binding prot
8	1350.5	77.3	303	1 A38274	B box-binding prot
9	1304.5	74.7	305	1 S22313	enhancer factor pr
10	1246	71.3	348	2 A49594	DNA-binding protei
11	840	48.1	372	2 S69501	RNA-binding protei
12	773	44.2	342	1 I53354	RYB-a protein - ra
13	742.5	42.5	291	2 S51608	RYB-a protein - ra
14	742.5	42.5	291	2 S48055	y box-binding prot
15	633	36.2	336	1 B38274	mRNA-binding prote
16	606.5	34.7	324	1 A41766	transcription enha
17	461	26.4	140	4 J02233	enhancer factor I
18	435	24.9	178	2 J02032	hypothetical prote
19	393	22.5	208	2 T21689	hypothetical prote
20	316.5	18.1	294	2 T19920	y box-binding prot
21	281	16.1	74	2 S68269	hypothetical prote
22	281	16.1	267	2 T25767	hypothetical prote
23	265.5	15.2	265	2 T29322	enhancer factor I
24	200.5	11.5	83	2 JC2021	glycine-rich prote
25	186.5	10.7	214	1 KNT25	glycine-rich prote
26	178.5	10.2	201	2 P84536	cold shock-like pr
27	172	9.8	69	2 A84967	cold-shock protein
28	170.5	9.8	66	2 A89853	cold shock-like pr
29	169	9.7	69	2 AH0579	

30	168.5	9.6	65	2 B84101	cold-shock protein
31	168.5	9.6	203	1 JQ1061	glycine-rich prote
32	166.5	9.5	65	2 H97267	cold shock protein
33	164	9.4	69	2 S49050	cold shock protein
34	164	9.4	69	2 AG0316	probable cold shoc
35	164	9.4	69	2 A10212	cold shock protein
36	164	9.4	69	2 B85562	cold shock protein
37	164	9.4	69	2 F90711	cold shock protein
38	164	9.4	70	2 D64840	cold shock protein
39	164	9.4	70	2 A30772	cold shock-like pr
40	164	9.4	70	2 E85634	homolog of Salmone
41	163	9.3	69	2 S43618	cold shock protein
42	163	9.3	69	2 E30945	cold shock protein
43	163	9.3	69	2 A30727	cold shock-like pr
44	163	9.3	69	2 H85793	cold shock protein
45	161.5	9.2	66	2 AF1681	cold shock protein

ALIGNMENTS

RESULT 1

I39382
Y box-binding protein 1 - human
N:Alternate names: DNA binding protein B; transcription enhancer factor EF1a; transcrip
C:Species: Homo sapiens (man)
C>Date: 16-Feb-1996 #sequence revision 16-Feb-1996 #text_change 22-Jun-1999
C:Accession: I39382; PS0015; A40498
R:Horwitz, E.M.; Maloney, K.A.; Lev, T.J.
J. Biol. Chem. 269, 14130-14139, 1994
A:Title: A human protein containing a 'cold shock' domain binds specifically to H-DNA u
A:Reference number: A54085; MUID:94245734; PMID:8188694
A:Accession: I39382
A:Molecule type: mRNA
A:Residues: 1-324 <RES>
A:Cross-references: GB:L28809; NID:9454151; PIDN:AAA20871.1; PID:9454152
R:Sakura, H.; Maekawa, T.; Imamoto, F.; Yasuda, K.; Ishii, S.
Gene 73, 499-507, 1988
A:Title: Two human genes isolated by a novel method encode DNA-binding proteins contain
A:Reference number: PS0014; MUID:89211987; PMID:2977358
A:Accession: PS0015
A:Molecule type: mRNA
A:Residues: 'EFGQPORALSSPTAAAGLVITPREPQLPQAPVITAT', 1-324 <SAK>
A:Cross-references: GB:M24070; NID:G181485; PIDN:AAA35750.1; PID:G181486
R:Didier, D.K.; Schiftenbauer, J.; Woulfe, S.L.; Zacheis, M.; Schwartz, B.D.
Proc. Natl. Acad. Sci. U.S.A. 85, 7322-7326, 1988
A:Title: Characterization of the cDNA encoding a protein binding to the major histocomp
A:Reference number: A40498; MUID:89017190; PMID:3174636
A:Molecule type: mRNA
A:Residues: 1-119,'E',121-313,'RSRG' <DID>
A:Cross-references: GB:J03827; NID:G340418; PIDN:AAA61308.1; PID:G340419
C:Genetics:
A:Gene: GDB:YB1
A:Cross-references: GDB:5577123; OMIM:154030.
A:Map position: lp34-1p34
C:Superfamily: Y box-binding protein 1; cold shock domain homology
C:Keywords: DNA binding; nucleus; transcription regulation
F:61-125/Domain: cold shock domain homology <CSD>

Query Match 100.0%; Score 1747; DB 1; Length 324;
Best Local Similarity 100.0%; Pred. No. 4.6e-99;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSSEATQPPAAPPAALSAADTKPGITGSGAGSGPGGLTSAAPAGGDKKVIATKVL	60
DB	1	MSSEATQPPAAPPAALSAADTKPGITGSGAGSGPGGLTSAAPAGGDKKVIATKVL	60
QY	61	GVTKWPNVNGYGFNRNDTKEDVFVHQTAKKNNPKYLRVSGDGETVEFVVEGEKA	120
DB	61	GVTKWPNVNGYGFNRNDTKEDVFVHQTAKKNNPKYLRVSGDGETVEFVVEGEKA	120
QY	121	EAANTVPGGVFVQGSKYAADRNHYRYPRRRPPRRYQNNYQNSGSEKNGESSAPEG	180

```
Db 121 EAAVNTGGVPGVGGKYAADRNHYRRYPRRRGPPRYQNYQNSGEGKESGSAPEG 180
Qy 181 QAQRRPYRRRRFPYMYRRPYGRRPOYSNPPVQGVMEGADNQAGEQGRPVQRQNYRG 240
Db 181 QAQRRPYRRRRFPYMYRRPYGRRPOYSNPPVQGVMEGADNQAGEQGRPVQRQNYRG 240
Qy 241 YRPRFRGPPRQPRQREDGNEEDKENQDGTQGGQPPQRRYRNFYRRRRRPNPKPDG 300
Db 241 YRPRFRGPPRQPRQREDGNEEDKENQDGTQGGQPPQRRYRNFYRRRRRPNPKPDG 300
Qy 301 KETKAADPPAENSSAPEAEOGGAE 324
Db 301 KETKAADPPAENSSAPEAEOGGAE 324

RESULT 2
Y box-binding protein 1 - bovine
N:Alternate names: DNA binding protein B; transcription enhancer factor Bf1a; transcript
C:Species: Bos primigenius taurus (cattle)
C>Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 22-Jun-1999
C:Accession: JQ2292
R:Ozer, J.; Chalkley, R.; Sealy, L.
Gene 124, 223-230, 1993
A:Title: Isolation of the CCAAT transcription factor subunit EF1A cDNA and a potentially
family.
A:Reference number: JQ2292; MUID:93185927; PMID:8444345
A:Accession: JQ2292
A:Molecule type: mRNA
A:Residues: 1-324 <OE>
A:Cross-references: GB:M95793; NID:g162982; PIDN:AAA30497.1; PID:g162983
C:Comment: This protein is a multi-subunit trans-acting complex that binds to CCAAT box
C:Superfamily: Y box-binding protein 1; cold shock domain homology
C:Keywords: DNA binding; nucleus; transcription regulation
F:61-125/Domain: cold shock domain homology <CSD>

Query Match 100.0%; Score 1747; DB 1; Length 324;
Best Local Similarity 100.0%; Pred. No. 4.6e-99;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSSEAEQQPPAAPALSAADTKPGTTGSGAGSGPGGLTSAAPAGDGKKVIATKVL 60
Db 1 MSSEAEQQPPAAPALSAADTKPGTTGSGAGSGPGGLTSAAPAGDGKKVIATKVL 60
Qy 61 GTVKWFWNRNGYGFNRNDTKEDVFVHQTAKNNPKYLSVGDGETVEFDVVEGEKA 120
Db 61 GTVKWFWNRNGYGFNRNDTKEDVFVHQTAKNNPKYLSVGDGETVEFDVVEGEKA 120
Qy 121 EAAVNTGGVPGVGGKYAADRNHYRRYPRRRGPPRYQNYQNSGEGKESGSAPEG 180
Db 121 EAAVNTGGVPGVGGKYAADRNHYRRYPRRRGPPRYQNYQNSGEGKESGSAPEG 180
Qy 181 QAQRRPYRRRRFPYMYRRPYGRRPOYSNPPVQGVMEGADNQAGEQGRPVQRQNYRG 240
Db 181 QAQRRPYRRRRFPYMYRRPYGRRPOYSNPPVQGVMEGADNQAGEQGRPVQRQNYRG 240
Qy 241 YRPRFRGPPRQPRQREDGNEEDKENQDGTQGGQPPQRRYRNFYRRRRRPNPKPDG 300
Db 241 YRPRFRGPPRQPRQREDGNEEDKENQDGTQGGQPPQRRYRNFYRRRRRPNPKPDG 300
Qy 301 KETKAADPPAENSSAPEAEOGGAE 324
Db 301 KETKAADPPAENSSAPEAEOGGAE 324

RESULT 3
A55971
Y box-binding protein 1 - rabbit
N:Alternate names: DNA binding protein B; transcription enhancer factor Bf1a; transcript
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 22-Jun-1999
C:Accession: A55971
```

```
R:Evdokimova, V.M.; Wei, C.L.; Sitikov, A.S.; Simonenko, P.N.; Lazarev, O.A.; Vasilenko,
J. Biol. Chem. 270, 3186-3192, 1995
A:Title: The major protein of messenger ribonucleoprotein particles in somatic cells is
A:Reference number: A55971; MUID:95155408; PMID:7852402
A:Accession: A55971
A:Molecule type: mRNA
A:Residues: 1-324 <EVD>
A:Cross-references: GB:U16821; NID:g608517; PIDN:AAA6069.1; PID:g608518
C:Superfamily: Y box-binding protein 1; cold shock domain homology
C:Keywords: DNA binding; phosphoprotein; RNA binding; transcription regulation; translat
F:61-125/Domain: cold shock domain homology <CSD>

Query Match 98.9%; Score 1728; DB 1; Length 324;
Best Local Similarity 98.8%; Pred. No. 6.5e-98;
Matches 320; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSSEAEQQPPAAPALSAADTKPGTTGSGAGSGPGGLTSAAPAGDGKKVIATKVL 60
Db 1 MSSEAEQQPPAAPALSAADTKPGTTGSGAGSGPGGLTSAAPAGDGKKVIATKVL 60
Qy 61 GTVKWFWNRNGYGFNRNDTKEDVFVHQTAKNNPKYLSVGDGETVEFDVVEGEKA 120
Db 61 GTVKWFWNRNGYGFNRNDTKEDVFVHQTAKNNPKYLSVGDGETVEFDVVEGEKA 120
Qy 121 EAAVNTGGVPGVGGKYAADRNHYRRYPRRRGPPRYQNYQNSGEGKESGSAPEG 180
Db 121 EAAVNTGGVPGVGGKYAADRNHYRRYPRRRGPPRYQNYQNSGEGKESGSAPEG 180
Qy 181 QAQRRPYRRRRFPYMYRRPYGRRPOYSNPPVQGVMEGADNQAGEQGRPVQRQNYRG 240
Db 181 QAQRRPYRRRRFPYMYRRPYGRRPOYSNPPVQGVMEGADNQAGEQGRPVQRQNYRG 240
Qy 241 YRPRFRGPPRQPRQREDGNEEDKENQDGTQGGQPPQRRYRNFYRRRRRPNPKPDG 300
Db 241 YRPRFRGPPRQPRQREDGNEEDKENQDGTQGGQPPQRRYRNFYRRRRRPNPKPDG 300
Qy 301 KETKAADPPAENSSAPEAEOGGAE 324
Db 301 KETKAADPPAENSSAPEAEOGGAE 324

RESULT 4
I58195
Y box-binding protein 1 - mouse
N:Alternate names: DNA binding protein B; MSY1 protein; transcription enhancer factor EF
C:Species: Mus musculus (house mouse)
C>Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 22-Jun-1999
C:Accession: I58195; A45976; I52910; S22822
R:Wolffe, A.P.; Tataru, S.; Ranjan, M.; Familari, M.
New Biol. 4, 290-298, 1992
A:Title: The Y-box factors: a family of nucleic acid binding proteins conserved from Esc
A:Reference number: I58195; MUID:92322631; PMID:1622927
A:Accession: I58195
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-322 <RS>
A:Cross-references: GB:M62867; NID:g199820; PIDN:AAA63390.1; PID:g199821
R:Tatari, S.R.; Familari, M.; Wolffe, A.P.
J. Biol. Chem. 268, 12213-12220, 1993
A:Title: A mouse Y box protein, MSY1, is associated with paternal mRNA in spermatocytes.
A:Reference number: A45976; MUID:93280200; PMID:8505341
A:Accession: A45976
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-322 <TAR>
R:Shaughnessy, M.; Lee, D.; Wistow, G.J.
Curr. Eye Res. 11, 171-181, 1992
A:Title: Absence of MHC expression in lens and cloning of dbpB/YB-1, a DNA-binding prote
A:Reference number: I52910
A:Accession: I52910
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-28, 'G', 30-322 <RE2>
```


A;Cross-references: GB:M60419; NID:G202434; PIDN:AAA40577.1; PID:G202435
A;Experimental source: lens
R;Gal, X.; Lipson, K.E.; Prystowsky, M.B.
Nucleic Acids Res. 20, 601-606, 1992
A;Title: Unusual DNA binding characteristics of an in vitro translation product of the C
A;Reference number: S22822; MUID:92158671; PMID:1741293
A;Accession: S22822
A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL
A;Molecule type: mRNA
A;Residues: 1-28, 'G', '30-42', 'RR', '46-237', 'P', '239-322 <AI>
A;Cross-references: EMBL:X57621; NID:G95450; PIDN:CAA40847.1; PID:G55451
A;Experimental source: strain C57BL/6
C;Genetics:
A;Gene: MSY-1; YB-1
C;Superfamily: Y box-binding protein 1; cold shock domain homology
C;Keywords: DNA binding; eye lens; nucleus; RNA binding; testis; transcription regulation
F;59-123/Domain: cold shock domain homology <CSD>

Query Match 97.4%; Score 1701; DB 1; Length 322;
Best Local Similarity 98.1%; Pred. No. 2.8e-96;
Matches 318; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 1 MSSEATQPPAAPALSAADTKPGTTGGAGSGGPGGLTSAAPAGDKKVIATKVL 60
Db 1 MSSEATQPPAAP -AAALSAADTKPGTASGAGSGGPGGLTSAAPAGDKKVIATKVL 58
QY 61 GTVKNFNVANGVGFINRNTDKEDVFHQTAIKKNPRKYLRSVGDGETVEFDDVEGEKA 120
Db 59 GTVKNFNVANGVGFINRNTDKEDVFHQTAIKKNPRKYLRSVGDGETVEFDDVEGEKA 118
QY 121 EAANVTGPGVPVQGSKYAADRNHYRYPRRRPPPNYOQYNSSEBKNKGSSAPEG 180
Db 119 EAANVTGPGVPVQGSKYAADRNHYRYPRRRPPPNYOQYNSSEBKNKGSSAPEG 178
QY 181 QACQRPYRRRPPPYMYRRPYGRRPOYNNPVOGMEGADNQGAGEGRPVQRNNYRG 240
Db 179 QACQRPYRRRPPPYMYRRPYARRPOYNNPVOGMEGADNQGAGEGRPVQRNNYRG 238
QY 241 YRPRFRGPRQRPREDGNEEDKENQGDTCQCPQRRYRNFYRRRPNKPDG 300
Db 239 YRPRFRGPRQRPREDGNEEDKENQGDTCQCPQRRYRNFYRRRPNKPDG 298
QY 301 KETKAADPPAENSAPAEQGGAE 324
Db 299 KETKAADPPAENSAPAEQGGAE 322

RESULT 5
A23677
Y box-binding protein 1 - rat
N;Alternate names: DNA binding protein B; transcription enhancer factor Bfla; transcript
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Oct-1991 #sequence_revision 31-Jul-1992 #text_change 22-Jun-1999
C;Accession: A23677
R;Ozer, J.; Faber, M.; Chalkley, R.; Sealy, L.
J. Biol. Chem. 285, 22143-22152, 1990
A;Title: Isolation and characterization of a cDNA clone for the CCAAT transcription factor
C;Superfamily: Y box-binding protein 1; cold shock domain homology
C;Keywords: DNA binding; nucleus; transcription regulation
F;59-123/Domain: cold shock domain homology <CSD>

Query Match 97.3%; Score 1699; DB 1; Length 322;
Best Local Similarity 98.1%; Pred. No. 3.7e-96;
Matches 318; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 1 MSSEATQPPAAPALSAADTKPGTTGGAGSGGPGGLTSAAPAGDKKVIATKVL 60
Db 1 MSSEATQPPAAP -AAALSAADTKPGTASGAGSGGPGGLTSAAPAGDKKVIATKVL 58

C:Species: Homo sapiens (man)
C:Date: 02-Dec-1993 #sequence_revision 26-May-1995 #text_change 07-May-1999
C:Accession: S34426
R:Kolluri, R.; Kinniburgh, A.J.
Nucleic Acids Res. 19, 4771, 1991
A:Title: Full length cDNA sequence encoding a nuclease-sensitive element DNA binding protein
A:Reference number: S34426; MUID:91367681; PMID:1891370
A:Accession: S34426
A>Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL
A:Molecule type: mRNA
A:Residues: 1-321 <KOL>
A:Cross-references: EMBL:M85234
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1992
C:Genetics:
A:Gene: GDB:NSEPI
A:Cross-references: GDB:128796
C:Superfamily: Y box-binding protein 1; cold shock domain homology
C:Keywords: DNA binding; nucleus; transcription regulation
F:60-124/Domain: cold shock domain homology <CSD>

Query Match 79.2%; Score 1383; DB 1; Length 321;
Best Local Similarity 82.1%; Pred. No. 5.1e-77;
Matches 271; Conservative 6; Mismatches 37; Indels 16; Gaps 3;

QY 1 MSSEAEQTQPPAAPALSAADTKPGTTGSGAGS-----GGPGGLTSAAPAGGDKK 53
DB 1 MSSEAEQTQPPAAPALSAADTKPGTTGAGAVARAHHGGAG-----AGDKK 52
QY 54 VIATKVLGTWKFNVRNGYGFINRNDTKEDVHVHTAIKKNPRKYLRSVGDGETVEFDV 113
DB 53 VIATKVLGTWKFNVRNGYGFINRNDTKEDVHVHTAIKKNPRKYLRSVGDGETVEFDV 112
QY 114 VEGSKGAEEANVTGPGVPGVQSKYAADRHHYRRYPRRRPQYNNPPVQGEVMEGADNQAGEQGRPV 173
DB 113 VEGSKGAEEANVTGPGVPGVQSKYAADRHHYRRYPRRRPQYNNPPVQGEVMEGADNQAGEQGRPV 172
QY 174 SESAPQQAQOORPYRRRRPPYRRYPRRRPQYNNPPVQGEVMEGADNQAGEQGRPV 233
DB 173 SESAPQQAQOORPYRRRRPPYRRYPRRRPQYNNPPVQGEVMEGADNQAGEQGRPV 231
QY 234 RQNYVGRVPRFRGPRQPRQEDGNEEDKENQDGTQCGQPPQRRYRRNFRNRRRPE 293
DB 232 RQICIGDIDHDSAGALLAKQPRQEDGNEEDKENQDGTQCGQPPQRRYRRNFRNRRRPE 291
QY 294 NPKPDQKTKAADPPAENSAPAEQGG 323
DB 292 NPKPDQKTKAADPPAENSAPAEQGG 321

RESULT 8
A38274
Y box-binding protein 1 - African clawed frog
N:Alternate names: DNA binding protein B; transcription enhancer factor Bf1a; transcript
C:Species: Xenopus laevis (African clawed frog)
C:Date: 21-Jun-1991 #sequence_revision 21-Jun-1991 #text_change 22-Jun-1999
C:Accession: A38274
R:Tatari, S.R.; Wolffe, A.P.
Proc. Natl. Acad. Sci. U.S.A. 87, 9028-9032, 1990
A:Title: Xenopus Y-box transcription factors: molecular cloning, functional analysis, and
A:Reference number: A38274; MUID:91062413; PMID:2247479
A:Accession: A38274
A:Molecule type: mRNA
A:Residues: 1-303 <TAF>
A:Cross-references: GB:M59453; GB:M38382; NID:g214154; PID:AA49715.1; PID:g214155
C:Superfamily: Y box-binding protein 1; cold shock domain homology
C:Keywords: DNA binding; nucleus; transcription regulation
F:39-103/Domain: cold shock domain homology <CSD>

Query Match 77.3%; Score 1350.5; DB 1; Length 303;
Best Local Similarity 79.7%; Pred. No. 4.5e-75;
Matches 259; Conservative 9; Mismatches 34; Indels 23; Gaps 3;

QY 1 MSSEAEQTQPPAAPALSAADTKPGTTGSGAGSGPGGLTSAAPAGGDKKVIATKVL 60

DB 1 MSSEVETQ-----QQPDALGKAG-----BPAATVGGKKVIATKVL 38
QY 61 GTVKFNVRNGYGFINRNDTKEDVHVHTAIKKNPRKYLRSVGDGETVEFDVVEGEKGA 120
DB 39 GTVKFNVRNGYGFINRNDTKEDVHVHTAIKKNPRKYLRSVGDGETVEFDVVEGEKGA 98
QY 121 EAANTVGTGPGVPGVQSKYAADRHHYRRYPRRRPQYNNPPVQGEVMEGADNQAGEQGRPV 180
DB 99 EAANTVGTGPGVPGVQSKYAADRHHYRRYPRRRPQYNNPPVQGEVMEGADNQAGEQGRPV 158
QY 181 -CAQORRYPYRRRRPQYNNPPVQGEVMEGADNQAGEQGRPV 239
DB 159 DSNQORPYRRRRPQYNNPPVQGEVMEGADNQAGEQGRPV 218
QY 240 GYRFRFRGPRQPRQEDGNEEDKENQDGTQCGQPPQRRYRRNFRNRRRPPQD 299
DB 219 GYRFRFRGPRQPRQEDGNEEDKENQDGTQCGQPPQRRYRRNFRNRRRPPQD 278
QY 300 GKTKAADPPAENSAPAEQGG 324
DB 279 GKTKAADPPAENSAPAEQGG 303

Query Match 74.7%; Score 1304.5; DB 1; Length 305;
Best Local Similarity 78.4%; Pred. No. 2.8e-72;
Matches 258; Conservative 13; Mismatches 29; Indels 29; Gaps 6;

QY 1 MSSEAEQTQPPAAPALSAADTKPGTTGSGAGSGPGGLTSAAPAGGDKKVIATK 58
DB 1 MSSEVETQEQPDALG-----GKAGQEPATV-GEKKVIATK 36
QY 59 VLGTVKFNVRNGYGFINRNDTKEDVHVHTAIKKNPRKYLRSVGDGETVEFDVVEGEK 118
DB 37 VLGTVKFNVRNGYGFINRNDTKEDVHVHTAIKKNPRKYLRSVGDGETVEFDVVEGEK 96
QY 119 GAEANVTGPGVPGVQSKYAADRHHYRRYPRRRPQYNNPPVQGEVMEGADNQAGEQGRPV 178
DB 97 GAEANVTGPGVPGVQSKYAADRHHYRRYPRRRPQYNNPPVQGEVMEGADNQAGEQGRPV 156
QY 179 EQQ-AQORRYPYRRRRPQYNNPPVQGEVMEGADNQAGEQGRPV 236
DB 157 EGEQNRQORPCRRRRPQYNNPPVQGEVMEGADNQAGEQGRPV 216
QY 237 MYRGRFRFRGPRQPRQEDGNEEDKENQDGTQCGQPPQRRYRRNFRNRRRPP 295
DB 217 MYRGRFRFRGPRQPRQEDGNEEDKENQDGTQCGQPPQRRYRRNFRNRRRPP 276
QY 296 KPDQKTKAADPPAENSAPAEQGG 324

RESULT 9
S22313
B box-binding protein - African clawed frog
N:Alternate names: YB3 protein
C:Species: Xenopus laevis (African clawed frog)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 22-Jun-1999
C:Accession: S22313
R:Cohen, I.; Reynolds, W.F.
Nucleic Acids Res. 19, 4753-4759, 1991
A:Title: The Xenopus YB3 protein binds the B box element of the class III promoter.
A:Reference number: S22313; MUID:91367675; PMID:1891365
A:Accession: S22313
A:Molecule type: mRNA
A:Residues: 1-305 <COH>
A:Cross-references: EMBL:X60217; NID:g65270; PID:CAA42778.1; PID:g65271
A:Experimental source: oocyte
C:Function:
A:Description: binds to the B-box promoter element; may be involved in class III gene re

Db 277 KPQGGKETKAAETSAENTSAPAEQGGAE 305

RESULT 10

A49594

enhancer factor protein 1 - chicken (fragment)

C:Species: Gallus gallus (chicken)

C>Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 20-Aug-1999

C:Accession: A49594

R:Kandala, J.C.; Guntaka, R.V.

Virolgy 198, 514-523, 1994

A:Title: Cloning of Rous sarcoma virus enhancer factor genes. I. Evidence that RSV-EF 1

A:Reference number: A49594; MUID:94120725; PMID:8291233

A:Keywords: DNA binding; nucleus

F:93-157/Domain: cold shock domain homology <CSD>

Query Match 71.3%; Score 1246; DB 2; Length 348;

Best Local Similarity 75.0%; Pred. No. 1.1e-68;

Matches 243; Conservative 15; Mismatches 48; Indels 18; Gaps 4;

QY 5 AETQCP-----PAAPPAAPALSAADTKPGTTGSGAGSGGPGGILTSAPAGGDKKIATK 58

Db 4 AETQPPAAPVPAPPPPTPNLTAA-----AGTAAAWPRRLRRGQEGH-----RNE 52

QY 59 VLGTVKWFNVRNGYGFNRNDTKEDVFVHQTAIKKNNPKYLSRVSVDGTEVDFDVEGK 118

Db 53 GFGNSEMVNVRNGYGLNRNDTKEDVFVHQTAIKKNNPKYLSRVSVDGTEVDFDVEGK 112

QY 119 GAERANVTGPGVQVQSKYAADNHNHVRYPVRRGPRPNVQYQNSGSEKNGESAP 178

Db 113 GAERANVTGPGVQVQSKYAADNHNHVRYPVRRGPRPNVQYQNSGSEKNGEAGNIP 172

QY 179 EGQAQQRPPYRRPPYRRPPYRRPPYRRPPYRRPPYRRPPYRRPPYRRPPYRRPPY 238

Db 173 EGQAQQRPPYRRPPYRRPPYRRPPYRRPPYRRPPYRRPPYRRPPYRRPPYRRPPY 231

QY 239 RGYRFRFRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPR 298

Db 232 RGYRFRFRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPR 291

QY 299 DGKETKAADPPAENSSAPAEQGG 322

Db 292 DGKETKAADPPAENSSAPAEQAG 315

RESULT 11

S69501

DNA-binding protein A variant - human

N:Alternate names: cold shock domain protein A

C:Species: Homo sapiens (man)

C>Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000

C:Accession: S69501

R:Colles, L.S.; Diamond, P.; Occhiodoro, F.; Vadas, M.A.; Shannon, M.F.

Nucleic Acids Res. 24, 2311-2317, 1996

A:Title: Cold shock domain proteins repress transcription from the GM-CSF promoter.

A:Reference number: S69501; MUID:96279731; PMID:8710501

A:Accession: S69501

A:Status: preliminary

A:Keywords: nucleic acid sequence not shown; translation not shown

F:93-157/Domain: cold shock domain homology <CSD>

Query Match 48.1%; Score 840; DB 2; Length 372;

Best Local Similarity 52.3%; Pred. No. 4.6e-44;

Matches 202; Conservative 25; Mismatches 83; Indels 76; Gaps 17;

QY 1 MSSEAE-----TQPPPAAPPAALSAADTKP-GTTGSGAG-----SGGPGGLT 43

Db 1 MSEAGEATTTTTTLTLPQAPTEAAAAAQQDPAPKPSVSGAPQAAAPAAPAAHVAGNPGG-- 58

QY 44 SAAPAG-----GDKVIATKVLGTGVKFNVRNGYGFNRNDTKEDVFV 86

Db 59 DAAPPAATGTAAASLATAAGSEDAEKVLAATKVLGTGVKFNVRNGYGFNRNDTKEDVFV 118

QY 87 HQTATKKNPKYLSRVSVDGTEVDFDVEGKGAEEAANVTGPGVQVQSKYAADNHNH 146

Db 119 HQTATKKNPKYLSRVSVDGTEVDFDVEGKGAEEAANVTGPGVQVQSKYAADNHNH 178

QY 147 R--YRRRGPRPNVQYQNSGSEKNGESAPEG-----QAQQRPP-- 187

Db 179 RYVYRRRGPRPNVQYQNSGSEKNGESAPEG-----QAQQRPP-- 230

QY 188 -YRRRRFPYRRPPYRRPPYRRPPYRRPPYRRPPYRRPPYRRPPYRRPPYRRPPY 241

Db 231 QYRQRFPYRRPPYRRPPYRRPPYRRPPYRRPPYRRPPYRRPPYRRPPYRRPPY 287

QY 242 RRRFR-RGPRRQPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPR 298

Db 288 RRRFR-RGPRRQPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPRGPR 347

QY 299 DGKETKAADPPAENSSAPAEQGGAE 324

Db 348 DGKETKAADPPAENSSAPAEQGGAE 372

RESULT 12

I53354

DNA-binding protein A - human

C:Species: Homo sapiens (man)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Apr-2000

C:Accession: I53354; PS0014; S65945

R:Kudo, S.; Mattei, M.G.; Fukuda, M.

Eur. J. Biochem. 231, 72-82, 1995

A:Title: Characterization of the gene for dbpA, a family member of the nucleic-acid-binding protein A.

A:Reference number: I53354; MUID:95354705; PMID:7628487

A:Accession: I53354

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-342 <KUD>

A:Cross-references: GB:L29071; NID:G950337; PIDN:AAA79243.1; PID:G950340

A:Experimental source: placenta

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1995

R:Sakura, H.; Maekawa, T.; Imamoto, F.; Yasuda, K.; Ishii, S.

Gene 73, 499-507, 1988

A:Title: Two human genes isolated by a novel method encode DNA-binding proteins contain

A:Reference number: PS0014; MUID:89211987; PMID:2977358

A:Accession: PS0014

A:Molecule type: mRNA

A:Residues: EFGRGSPRRARRSSRLRQDPTSAAGLRRIIRPGLPESEPRPPPPAALTADQPPRRRLSRRGGG', 1-

A:Cross-references: GB:M24069; NID:G181483; PIDN:AAA35749.1; PID:G181484

C:Genetics:

A:Gene: GDB:CSDA; dbpA

A:Cross-references: GDB:9865772; OMIM:603437

A:Map position: 12p13.1-12p13.1

A:Introns: 88/1; 109/2; 120/3; 150/3; 191/3; 260/3; 293/2

C:Superfamily: Y box-binding protein 1; cold shock domain homology

C:Keywords: DNA binding; nucleus; transcription regulation

F:93-157/Domain: cold shock domain homology <CSD>

Query Match 44.2%; Score 773; DB 1; Length 342;

Best Local Similarity 52.1%; Pred. No. 4.9e-40;

Matches 185; Conservative 24; Mismatches 72; Indels 74; Gaps 15;

QY 1 MSSEAE-----TQPPPAAPPAALSAADTKP-GTTGSGAG-----SGGPGGLT 43

Db 1 MSEAGEATTTTTTLTLPQAPTEAAAAAQQDPAPKPSVSGAPQAAAPAAPAAHVAGNPGG-- 58

44 SAAPAG-----GDKVIATKVLGTVKFNVRNGYGFINRNDTKEDVVF 86
 59 DAAPATGTAASAAAGSADAEEKVLATKVLGTVKFNVRNGYGFINRNDTKEDVVF 118
 87 HQTAKKNNPKYLSVGDGETVBDVVEGKGAEAAANVTGGVPGVQGSKYAADRNHYR 146
 119 HQTAKKNNPKYLSVGDGETVBDVVEGKGAEAAANVTGGVPGVQGSKYAADRNHYR 178
 147 R--YPRRRGPPRNYQNTQNSSEKNGESAPAG-----QAQORRP--- 187
 179 RGYVRRGPPRNY-----AGEEBEGSGSGEDFPATDRQFSGARNQLRPPQYRP 230
 188 -YRRRRFPYVRRPYGRPO---YSNPFGVQGVVEGADN--QAGBQGRPVQNNVYRGY 241
 231 QYRQRRFPYVHGQTFDRSRVLPHPNRQAGEIGEMKMDGVPEGAQLQG-PVHRN--PTY 287
 242 RPRFR-RGPPRQRPREDGNEEDKENQDETQGOQPPQRR-YRNNFNVRRRPEN 294
 288 RPRRSRGPFRPAPAVGEAEDKENQATSGPNQPSVRRGYRRPYNRRPPSS 342

RESULT 13
 S51608
 RYB-a protein - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 20-Jun-2000
 C:Accession: S51608
 R:Rito, K.; Tsutsumi, K.I.; Kuzumaki, T.; Gomez, P.F.; Otsu, K.; Ishikawa, K.
 submitted to the EMBL Data Library, February 1994
 A:Reference number: S51608
 A:Accession: S51608
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-291 <I>
 A:Cross-references: EMBL:D28557; NID:G505133; PID:G505133
 C:Superfamily: Y box-binding protein 1; cold shock domain homology
 F:84-148/Domain: cold shock domain homology <CSD>

Query Match 42.5%; Score 742.5; DB 2; Length 291;
 Best Local Similarity 51.0%; Pred. No. 2.9e-38;
 Matches 171; Conservative 16; Mismatches 75; Indels 73; Gaps 9;

QY 2 SSEATQPPAAPPAALSAADTKPGTTGSGAGS-----GGPGGLTSAAPAGDKKV 54
 DB 18 AADAPAAAPPDPAPKSPAAAGAPAPAPALLAGAPARRSPRAPGLISPRGKREKKV 77
 QY 55 IATKVLGTVKFNVRNGYGFINRNDTKEDVVFHQTAKKNNPKYLSVGDGETVBDVVF 114
 DB 78 LATKVLGTVKFNVRNGYGFINRNDTKEDVVFHQTAKKNNPKYLSVGDGETVBDVVF 137
 QY 115 EGEKGAEEAANVTGGVPGVQGSKYAADRNHYR--YPRRRGPPRNYQNTQNSSEKNE 172
 DB 138 EGEKGAEEAANVTGGVPGVQGSKYAADRNHYR--YPRRRGPPRNYQNTQNSSEKNE 188
 QY 173 GSSEAPG-QAQRPPYRRRRFPYVRRPYGRPPQYNSNPPVQGVMEGADNQAGEOGR 231
 DB 189 MKDGVPEGAQLQVHR-----NPT----- 206
 QY 232 PVRQNNYRGYRPRFRGPPRQRPREDGNEEDKENQDETQGOQPPQRR-YRNNFNHYR 290
 DB 207 -----YRPRFRGPPRPAIPAIGEAEDKENQAANGPNQPSARRGFRFPYNYRR 257
 QY 291 -RPNPKPDQKTKAADPPAENSSAPEAQGGAE 324
 DB 258 PRPLNAVSDQKTKAGEAPTEN-PAPATEQSSAE 291

RESULT 14
 S48055
 RYB-a protein - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 26-Feb-1998
 C:Accession: S48055

R:Rito, K.; Tsutsumi, K.; Kuzumaki, T.; Gomez, P.F.; Otsu, K.; Ishikawa, K.
 Nucleic Acids Res. 22, 2036-2041, 1994
 A:Title: A novel growth-inducible gene that encodes a protein with a conserved cold-shock
 A:Reference number: S48055; MUID:94301785; PMID:8029009
 A:Accession: S48055
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-291 <I>
 A:Cross-references: EMBL:D28557
 C:Superfamily: Y box-binding protein 1; cold shock domain homology
 F:84-148/Domain: cold shock domain homology <CSD>

Query Match 42.5%; Score 742.5; DB 2; Length 291;
 Best Local Similarity 51.0%; Pred. No. 2.9e-38;
 Matches 171; Conservative 16; Mismatches 75; Indels 73; Gaps 9;

QY 2 SSEATQPPAAPPAALSAADTKPGTTGSGAGS-----GGPGGLTSAAPAGDKKV 54
 DB 18 AADAPAAAPPDPAPKSPAAAGAPAPAPALLAGAPARRSPRAPGLISPRGKREKKV 77
 QY 55 IATKVLGTVKFNVRNGYGFINRNDTKEDVVFHQTAKKNNPKYLSVGDGETVBDVVF 114
 DB 78 LATKVLGTVKFNVRNGYGFINRNDTKEDVVFHQTAKKNNPKYLSVGDGETVBDVVF 137
 QY 115 EGEKGAEEAANVTGGVPGVQGSKYAADRNHYR--YPRRRGPPRNYQNTQNSSEKNE 172
 DB 138 EGEKGAEEAANVTGGVPGVQGSKYAADRNHYR--YPRRRGPPRNYQNTQNSSEKNE 188
 QY 173 GSSEAPG-QAQRPPYRRRRFPYVRRPYGRPPQYNSNPPVQGVMEGADNQAGEOGR 231
 DB 189 MKDGVPEGAQLQVHR-----NPT----- 206
 QY 232 PVRQNNYRGYRPRFRGPPRQRPREDGNEEDKENQDETQGOQPPQRR-YRNNFNHYR 290
 DB 207 -----YRPRFRGPPRPAIPAIGEAEDKENQAANGPNQPSARRGFRFPYNYRR 257
 QY 291 -RPNPKPDQKTKAADPPAENSSAPEAQGGAE 324
 DB 258 PRPLNAVSDQKTKAGEAPTEN-PAPATEQSSAE 291

RESULT 15
 R38274
 Y box-binding protein 2 - African clawed frog
 N:Alternate names: cytoplasmic mRNA-binding protein p54
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 21-Jun-1991 #sequence_revision 21-Jun-1991 #text_change 22-Jun-1999
 C:Accession: R38274; B41786
 R:Tatari, S.R.; Wolffe, A.P.
 Proc. Natl. Acad. Sci. U.S.A. 87, 9028-9032, 1990
 A:Title: Xenopus Y-box transcription factors: molecular cloning, functional analysis, and
 A:Reference number: A38274; MUID:91062413; PMID:2247479
 A:Accession: R38274
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-336 <I>
 A:Cross-references: GB:M59454; GB:M38382; NID:G214156; PID:G214157; GB:
 R:Murray, M.T.; Schiller, D.L.; Franke, W.W.
 Proc. Natl. Acad. Sci. U.S.A. 89, 11-15, 1992
 A:Title: Sequence analysis of cytoplasmic mRNA-binding proteins of Xenopus oocytes ident
 A:Reference number: A41786; MUID:92107999; PMID:1729676
 A:Accession: B41786
 A:Molecule type: protein
 A:Residues: 56-82,95-105, 'XXX', 109-113,234-253, 'A', 255 <I>
 C:Superfamily: Xenopus Y box-binding protein 2; cold shock domain homology
 C:Keywords: DNA binding; nucleus; oocyte; RNA binding; testis; transcription regulation
 F:44-108/Domain: cold shock domain homology <CSD>

Query Match 36.2%; Score 633; DB 1; Length 336;
 Best Local Similarity 41.8%; Pred. No. 1.5e-31;
 Matches 152; Conservative 43; Mismatches 91; Indels 78; Gaps 11;
 QY 3 SEATQPPAAPPAALSAADTKPGTTGSGAGSGPGGLTSAAPAGDKKVATKVLGT 62

Db	2	SEAEQEPFVP-QPESEPIQPG-----IAAARNQANKVLATQVGT	45
Qy	63	VKWFNVRNGYGFNRNDTKEDVHQTAKNNPRKYLRSVGDGETVEFDVVEGKGAEA	122
Db	46	VKWFNVRNGYGFNRNDTKEDVHQTAKNNPRKFLRSVGDGETVEFDVVEGKGAEA	105
Qy	123	ANYTGGGVPOGSKVAADRNNHYRR---YPRRRGPPRNYCQNYQNSGEGKNEGSESAPF	179
Db	106	ANYTGGGVPOGSKVAADRNNHYRR---YPRRRGPPRNYCQNYQNSGEGKNEGSESAPF	165
Qy	180	GQAQORRRPYRRRRPPYMYRRPYGRAPQYNSPPVQG-EVMEGADNQ-----GAGE	228
Db	166	QRPQRRRP-----PPFFYRRRRRGRFPNNQNOGAEVTEQSENKDPVAPTSEALASGD	219
Qy	229	Q-----GRPVQNMRYGRYRFRFRGPPRQOPREDGNEEDKENQDGTQOQOPPPQRRYR	283
Db	220	DPORPPRRRQRRFRFRFR-----PAQQTTEGGDGETKAESGEDPRPE--PQQRNR	272
Qy	284	NFNYYRRR-----PENPKPDGK-ETKAADPPAENS	313
Db	273	PYQRRRRQGATQVAATAQEGKAEPQHPASEGTPSDSPTDDGAPVQSSAPDPGIADT	332
Qy	314	SAPF	317
Db	333	PAPE	336

Search completed: August 25, 2004, 00:27:03
Job time : 42 secs

This Page Blank (uspo)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 00:07:24 ; Search time 25 Seconds
(without alignments)
674.829 Million cell updates/sec

Title: US-10-028-415-40

Perfect score: 1747

Sequence: 1 MSSEARTQPPAAPAPAL.....AADPPAENSSAPEAQGGAE 324

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1747	100.0	324	YB1_HUMAN	P16991 h nuclease
2	1707	97.7	322	YB1_MOUSE	P27817 m nuclease
3	1589.5	91.0	321	YB1_CHICK	Q06066 gallus gall
4	1350.5	77.3	303	YB1_XENLA	P21573 xenopus lae
5	1304.5	74.7	305	YB3_XENLA	Q00436 xenopus lae
6	865.5	49.5	361	DBPA_RAT	Q62764 rattus norv
7	840	48.1	372	DBPA_HUMAN	P16989 homo sapien
8	632	36.2	336	YB56_XENLA	P45441 xenopus lae
9	606.5	34.7	324	YB54_XENLA	P41824 aplysia cal
10	497.5	28.5	253	YBPH_APLCA	P27484 nicotiana s
11	186.5	10.7	214	GRP2_NICSY	Q38896 arabidopsis
12	178.5	10.2	201	GR2B_ARATH	P57560 buchnera ap
13	172	9.8	68	CSPE_BUCAL	Q89A90 buchnera ap
14	172	9.8	68	CSPE_BUCBP	Q83R19 shigella fl
15	167	9.6	68	CSPC_SHIFL	P36997 escherichia
16	164	9.4	68	CSPE_ECOLI	Q47110 escherichia
17	164	9.4	70	CSPG_ECOLI	P36996 escherichia
18	163	9.3	68	CSFC_ECOLI	P39818 salmonella
19	163	9.3	70	CSPJ_SALTY	Q45098 bacillus an
20	162.5	9.3	65	CSFC_BACLU	Q30875 micrococci
21	162	9.3	67	CSFA_MICLU	Q67327 aquifex aeo
22	161	9.2	70	CSPT_AQUAE	P39158 bacillus su
23	160.5	9.2	66	CSFC_BACSU	Q48770 listeria mo
24	159.5	9.1	66	CSPA_LISNO	P58726 salmonella
25	159	9.1	70	CSPJ_SALTY	P57407 buchnera ap
26	156	8.9	68	CSPE_BUCAL	Q89A90 buchnera ap
27	155.5	8.9	2716	CSA_DROME	Q81994 drosophila
28	154	8.8	70	CSPE_YERPE	Q85593 yersinia pe
29	152.5	8.7	67	CSPA_STRPY	Q34994 streptococc
30	152	8.7	71	CSPE_ECOLI	P36995 escherichia
31	151	8.6	68	CSFA_STIAU	P23265 stigmatella
32	150	8.6	69	CSFA_ECOLI	P52777 escherichia
33	149.5	8.6	963	YQ36_CABEL	Q09457 caenorhabdi

34	148.5	8.5	356	1	RS41_ARATH	P29266 arabidopsis
35	148	8.5	331	1	PRP1_HUMAN	P04280 homo sapien
36	147	8.4	70	1	CSPV_VIBCH	Q9K116 vibrio chol
37	146	8.4	67	1	CSPF_STRCO	P48859 streptomyce
38	146	8.4	70	1	CSPA_VIBCH	Q9K000 vibrio chol
39	146	8.4	70	1	CSFG_SHEVI	Q9L170 stewanella
40	145.5	8.3	66	1	CSFB_LISMO	P96791 listeria mo
41	145	8.3	1446	1	IE18_PVKKA	P33479 pseudorabie
42	144.5	8.3	67	1	CSPB_BACSU	P32081 bacillus su
43	143.5	8.2	65	1	CSPB_BACCE	Q45097 bacillus ce
44	143.5	8.2	66	1	CSPT_THEMEA	O54310 thermotoga
45	143.5	8.2	276	1	PRPL_HUMAN	P10162 homo sapien

ALIGNMENTS

RESULT 1
YB1_HUMAN
ID YB1_HUMAN STANDARD; PRT; 324 AA.
AC P16991; P16990;
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-AUG-1991 (Rel. 19, Last annotation update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Nuclease sensitive element binding protein 1 (Y box binding protein-1)
DE (Y-box transcription factor) (YB-1) (CCAAT-binding transcription
DE factor I subunit A) (CBF-A) (Enhancer factor I subunit A) (E1-A)
DE (DNA-binding protein B) (DBPB).
GN NSEPI OR YB1.
OS Homo sapiens (Human), and
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606, 9913;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Human;
RX MEDLINE=89211987; PubMed=2977358;
RA Sakura H., Maekawa T., Imamoto F., Yasuda K., Ishii S.;
RT "Two human genes isolated by a novel method encode DNA-binding
RT proteins containing a common region of homology."
Gene 73:499-507(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Human;
RX MEDLINE=89017190; PubMed=3174636;
RA Didier D.K., Schiffrinbauer J., Woulfe S.L., Zacheis M., Schwartz B.D.;
RT "Characterization of the cDNA encoding a protein binding to the major
RT histocompatibility complex class II Y box."
Proc. Natl. Acad. Sci. U.S.A. 85:7322-7326(1988).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=Human; TISSUE=Bone marrow;
RX MEDLINE=94245734; PubMed=818694;
RA Horwitz E.M., Maloney K.A., Ley T.J.;
RT "A human protein containing a 'cold shock' domain binds specifically
RT to H-DNA upstream from the human gamma-globin genes."
J. Biol. Chem. 269:14130-14139(1994).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=Human; TISSUE=Kidney, Muscle, and Uterus;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Wang J., Hsieh F.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gurnaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Mouse; TISSUE=Lung endothelial cells;
 RX MEDLINE=93185927; PubMed=8444345;
 RT Ozer J., Chalkley R., Sealy L.;
 RA "Isolation of the CCAAT transcription factor subunit EF1A cDNA and a
 RT potentially functional EF1A processed pseudogene from Bos taurus:
 RT insights into the evolution of the EF1A/dbpB/YB-1 gene family";
 RL Gene 124:223-230 (1993).
 CC -!- FUNCTION: Binds to CCAAT-containing Y box of HLA class II genes.
 CC Seems to be a negative regulatory factor.
 CC -!- SUBUNIT: This protein can bind to DNA as a homomeric form, (BFI-
 CC A)n or as a heteromeric form in association with BFI-B. Binds
 CC specifically to H-DNA upstream from the gamma-globin genes.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: Belongs to the cold-shock domain (CSD) family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M24070; AAA35750.1; ALT_INIT.
 DR EMBL; J03827; AAA61308.1; -.
 DR EMBL; L28809; AAA20871.1; -.
 DR EMBL; BC002411; AAH02411.1; -.
 DR EMBL; BC010430; AAH10430.1; -.
 DR EMBL; BC015208; AAH15208.1; -.
 DR EMBL; M95793; AAA30497.1; -.
 DR PIR; I39382; I39382.
 DR PIR; JQ2232; JQ2232.
 DR HSSP; P41016; IC90.
 DR TRANSFAC; T00186; -.
 DR TRANSFAC; T00910; -.
 DR Genew; HGNC:8014; NSEPI.
 DR GK; P16991; -.
 DR MIM; 154030; -.
 DR GO; GO:0005634; C:nucleus; NAS.
 DR GO; GO:0003677; F:DNA binding; NAS.
 DR GO; GO:0016564; F:transcriptional repressor activity; NAS.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
 DR InterPro; IPR002059; Cold shock.
 DR InterPro; IPR008994; Nucleic_acid_OB.
 DR Pfam; PF00313; CSD; 1.
 DR PRINTS; PR00050; COLDSHOCK.
 DR ProDom; PD00621; Cold_shock; 1.
 DR SMART; SM00357; CSP; 1.
 DR PROSITE; PS00352; COLD_SHOCK; 1.
 KW Transcription regulation; Repressor; DNA-binding; Nuclear protein.
 FT DOMAIN 61 125 CSD.
 FT CONFLICT 120 120 A -> E (IN REF. 2).
 FT CONFLICT 314 324 SAEAEQGGAE -> BSR (IN REF. 2).
 SQ SEQUENCE 324 AA; 35924 MW; DF0114BF974AED58 CRC64;
 Query Match 100.0%; Score 1747; DB 1; Length 324;
 Best Local Similarity 100.0%; Pred. No. 1.3e-100;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSSEAEFQQPPAAPALSAADTKGTGSGAGSGGGGLTSAPAGGDKKVIATKVL 60
 DB 1 MSSEAEFQQPPAAPALSAADTKGTGSGAGSGGGGLTSAPAGGDKKVIATKVL 60

QY 61 GTVKFNVNRYGFGFINRNDTKEDVFVHQTAIKKNPKRILSVGDETFEDVVEGEKGA 120
 DB 61 GTVKFNVNRYGFGFINRNDTKEDVFVHQTAIKKNPKRILSVGDETFEDVVEGEKGA 120
 QY 121 EAANTVGGGVVQSGKYAADRNHYRRYRRRGGPPRNYQNYQNSESEKKEGSESAPEG 180
 DB 121 EAANTVGGGVVQSGKYAADRNHYRRYRRRGGPPRNYQNYQNSESEKKEGSESAPEG 180
 QY 181 QAQQRPRYRRRPPYPYMYRRPYGRRPQYNSNPVQGVMEGADNQAGSQGRPVQNMVYG 240
 DB 181 QAQQRPRYRRRPPYPYMYRRPYGRRPQYNSNPVQGVMEGADNQAGSQGRPVQNMVYG 240
 QY 241 YPFRFRPPPRORPREDNEDKENQGDDETOGQOPPRYRRYRRRPPYRRRPPKPDG 300
 DB 241 YPFRFRPPPRORPREDNEDKENQGDDETOGQOPPRYRRYRRRPPYRRRPPKPDG 300
 QY 301 KETKAADPPAENSSAPEAEQGGAE 324
 DB 301 KETKAADPPAENSSAPEAEQGGAE 324
 RESULT 2
 YBI_MOUSE
 ID YBI_MOUSE STANDARD; PRT; 322 AA.
 AC P27817; P22569; P43482;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Nuclease sensitive element binding protein 1 (Y box binding protein-1)
 DE (Y-box transcription factor) (YB-1) (CCAAT-binding transcription
 DE factor I subunit A) (CBF-A) (Enhancer factor I subunit A) (EFI-A)
 DE (DNA-binding protein B) (DBPB).
 GN NSEPI OR YB1 OR YBX1 OR MSY-1.
 OS Mus musculus (Mouse), and
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090, 10116;
 RN [1] SEQUENCE FROM N.A.
 RC SPECIES=Mouse; TISSUE=Lens;
 RA Shaughnessy M., Wistow G.J.;
 RT "Absence of MHC expression in lens and cloning of dbpB/YB-1, a DNA-
 RT binding protein expressed in mouse lens.";
 RL Curr. Eye Res. 11:171-181 (1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Mouse;
 RX MEDLINE=9232631; PubMed=1622927;
 RA Familiari M., Sak D., Wolffe A.P., Tafuri S.R., Ranjan M.;
 RT "The Y-box factors: a family of nucleic acid binding proteins
 RT conserved from Escherichia coli to man.";
 RL New Biol. 4:290-298 (1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Mouse; STRAIN=C57BL/6;
 RX MEDLINE=92158671; PubMed=1741293;
 RA Gai X., Lipson K.E., Prystowsky M.B.;
 RT "Unusual DNA binding characteristics of an in vitro translation
 RT product of the CCAAT binding protein mYB-1.";
 RL Nucleic Acids Res. 20:601-606 (1992).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Mouse; STRAIN=FVB/N; TISSUE=Colon, and Mammary gland;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Boraldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blackesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Shailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[5]
RP SEQUENCE FROM N.A.
RC SPECIES=Homo; TISSUE=Liver;
RX MEDLINE=91093048; PubMed=1967130;
RA Ozer J., Faber M., Chalkley R., Sealy L.;
RT "Isolation and characterization of a cDNA clone for the CCAAT
RT transcription factor EF1A reveals a novel structural motif.";
RL J. Biol. Chem. 265:22143-22152(1990).
[6]
RN SEQUENCE FROM N.A.
RP SPECIES=Homo;
RC SPECIES=Homo;
RA Petty K.J., Bartalena L., Nikodem V.M.;
RL Submitted (XXX-1991) to the EMBL/GenBank/DBJ databases.
[7]
RN SEQUENCE FROM N.A.
RP SPECIES=Homo;
RC SPECIES=Homo;
RA Ogawa H., Date T., Nishizawa M., Pitot H.C., Fujioka M.;
RL Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Binds to CCAAT-containing Y box of HLA class II genes.
CC Seems to be a negative regulatory factor.
CC -!- SUBUNIT: This protein can bind to DNA as a homomeric form, (EF1-
CC A)n or as a heteromeric form in association with EF1-B.
CC -!- SUBCELLULAR LOCATION: Nuclear
CC -!- SIMILARITY: Belongs to the cold-shock domain (CSD) family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M60419; AAA40577.1; -
CC EMBL; M62867; AAA63390.1; -
CC EMBL; X57621; CAA40847.1; -
CC EMBL; BC013450; AAH13450.1; -
CC EMBL; BC013620; AAH13620.1; -
CC EMBL; M57299; AAA41108.1; -
CC EMBL; M69138; AAA40906.1; -
CC EMBL; D13309; BAA02569.1; -
CC PIR; A23677; A23677.
CC PIR; I58195; I58195.
CC HSSP; P41016; 1C90.
CC TRANSFAC; T00235; -
CC MGD; MGI:99146; Nsepl.
CC InterPro; IPR002059; Cold shock.
CC InterPro; IPR008994; Nucleic_acid_OB.
CC Pfam; PF00313; CSD; 1.
CC PRINTS; PR00050; COLDSHOCK.
CC ProDom; PD000621; Cold_shock; 1.
CC SMART; SM00357; CSP; 1.
CC PROSITE; PS00352; COLD_SHOCK; 1.
CC Transcription regulation; Repressor; DNA-binding; Nuclear protein.
FT DOMAIN 59 123 CSD
FT CONFLICT 5 5 A -> P (IN REF. 6).
FT CONFLICT 29 29 G -> A (IN REF. 2).
FT CONFLICT 43 45 AAP -> RR (IN REF. 3).
FT CONFLICT 238 238 G -> P (IN REF. 2).

FT CONFLICT 306 306 P -> S (IN REF. 5).
SQ SEQUENCE 322 AA; 35730 MW; C6667CF7CA10D45D CRC64;
Query Match 97.7%; Score 1707; DB 1; Length 322;
Best Local Similarity 98.5%; Pred. No. 3.8e-96;
Matches 319; Conservative 1; Mismatches 2; Indels 2; Gaps 1;
QY 1 MSSEAETQPPAAPPAAALSAADTKPTGTTGSGAGSGPGGLTSAAPAGDGKKVIATKVL 60
DB 1 MSSEAETQPPAAPP--AAALSAADTKPTGTTGSGAGSGPGGLTSAAPAGDGKKVIATKVL 58
QY 61 GTVKWFNVRNGYGFINRNDTKEDVFVHQTAKKNPKRYLSVGDGTVEFDVVEGEKGA 120
DB 59 GTVKWFNVRNGYGFINRNDTKEDVFVHQTAKKNPKRYLSVGDGTVEFDVVEGEKGA 118
QY 121 EAAVNTGPGVPGVQGSKYAADRNHYRRYPRRRGPPRNYQNYQNSSEKNEGSESAPEG 180
DB 119 EAAVNTGPGVPGVQGSKYAADRNHYRRYPRRRGPPRNYQNYQNSSEKNEGSESAPEG 178
QY 181 QAQRRPYRRRRFPYMYRRPYGRRPQYNSNPVQGVMEGADNQGAGEQGRPVQRNMYRG 240
DB 179 QAQRRPYRRRRFPYMYRRPYGRRPQYNSNPVQGVMEGADNQGAGEQGRPVQRNMYRG 238
QY 241 YRPRRRGPPRRORPREDGNEEDKENOGDETQGOQPPORRYRRNFRNRRRPPENPKPDG 300
DB 239 YRPRRRGPPRRORPREDGNEEDKENOGDETQGOQPPORRYRRNFRNRRRPPENPKPDG 298
QY 301 KETKAADPPAENSSAPEAEQGGAE 324
DB 299 KETKAADPPAENSSAPEAEQGGAE 322
RESULT 3
YBI_CHICK
ID YBI_CHICK STANDARD; PRT; 321 AA.
AC Q06066;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nuclease sensitive element binding protein 1 (Y box binding protein-1)
DE (Y-box transcription factor).
OS NSEPL OR YBI OR YB-1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93309452; PubMed=8321222;
RA Grant C.E., Deeley R.G.;
RT "Cloning and characterization of chicken YB-1: regulation of
RT expression in the liver.";
RL Mol. Cell. Biol. 13:4186-4196(1993).
CC -!- FUNCTION: Binds to CCAAT-containing Y box of HLA class II genes.
CC Seems to be a negative regulatory factor.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the cold-shock domain (CSD) family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L13032; AAA02573.1; -
CC PIR; A48136; A48136.
CC HSSP; P41016; 1C90.
CC TRANSFAC; T00236; -
CC InterPro; IPR002059; Cold shock.
CC InterPro; IPR008994; Nucleic_acid_OB.

DR Pfam; PF00313; CSD; 1.
 DR PRINTS; PR00050; COLD SHOCK.
 DR PRODOM; PR000621; Cold_shock; 1.
 DR SMART; SM00357; CSP; 1.
 DR PROSITE; PS00352; COLD_SHOCK; 1.
 DR Transcription regulation; Repressor; DNA-binding; Nuclear protein.
 FT DOMAIN 58 122 CSD.
 SQ SEQUENCE 321 AA; 35799 MW; 6496F306C1432274 CRC64;

Query Match 91.0%; Score 1589.5; DB 1; Length 321;
 Best Local Similarity 91.1%; Pred. No. 6e-91;
 Matches 296; Conservative 13; Mismatches 11; Indels 5; Gaps 3;

QY 1 MSSEAEETQPPAAP-PAAPALSAADTKPGTTGSGAGSGPGGLTSAAPAGGDKKVIATKV 59
 DB 1 MSSEAEET-QPPAAPVPAAPAAADSKP---NGSGNGSGSLASAAPAGGDKKVIATKV 56
 QY 60 LGTVKWFNVNGYGFINRNDTKEDVFVHQTAKKNNPKYLSRSGDGETVEFDVGEKG 119
 DB 57 LGTVKWFNVNGYGFINRNDTKEDVFVHQTAKKNNPKYLSRSGDGETVEFDVGEKG 116
 QY 120 AEAANVTGPGVPVQGSKYAADRNHYRYPERRGPPPNYQNNYQNSGKESGSAPE 179
 DB 117 AEAANVTGPGVPVQGSKYAADRNHYRYPERRGPPPNYQNNYQNSGKESGSAPE 176
 QY 180 GQAQRRPYRRRRPPPYMRRPYGRRPYQSNPPVQGEVMEGADNQGAGEQGRVPRQNM 239
 DB 177 GQAQRRPYRRRRPPPYMRRPYGRRPYQSNPPVQGEVMEGADNQGAGEQGRVPRQNM 236
 QY 240 GYRFRFRGPPRQPRQREDGNEEDKENQDGTQGGPPQRRYRNFYRRRRRPNKPD 299
 DB 237 GYRFRFRGPPRQPRQREDGNEEDKENQDGTQGGPPQRRYRNFYRRRRRPNKPD 296
 QY 300 GKETAADPPAENSSAPEAQGGAE 324
 DB 297 GKETAEPPEAENSSAPEAQGGAE 321

RESULT 4
 ID_YB1_XENLA STANDARD; PRT; 303 AA.
 AC P21573;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Nuclease sensitive element binding protein 1 (Y box binding protein-1)
 DE (Y-box transcription factor).
 GN NSEPI OR FRGY1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91062413; PubMed=2247479;
 RA Tafuri S.R., Wolffe A.P.;
 RT "Xenopus Y-box transcription factors: molecular cloning, functional
 analysis and developmental regulation.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:9028-9032(1990).
 CC -!- FUNCTION: BINDS TO CCAAT-CONTAINING Y BOX OF THE HSP70 GENES.
 CC SEEMS TO BE A NEGATIVE REGULATORY FACTOR.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: Belongs to the cold-shock domain (CSD) family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

DR EMBL; M59453; A38274; A38274.
 DR PIR; A38274; A38274.
 DR HSSP; P41016; IC90.
 DR TRANSFAC; T00293; -.
 DR InterPro; IPR002059; Cold_shock.
 DR InterPro; IPR008994; Nucleic_acid_OB.
 DR Pfam; PF00313; CSD; 1.
 DR PRINTS; PR00050; COLD SHOCK.
 DR PRODOM; PR000621; Cold_shock; 1.
 DR SMART; SM00357; CSP; 1.
 DR PROSITE; PS00352; COLD_SHOCK; 1.
 DR Transcription regulation; Repressor; DNA-binding; Nuclear protein.
 FT DOMAIN 39 103 CSD.
 SQ SEQUENCE 303 AA; 34633 MW; 691E8AB4216137C4 CRC64;

Query Match 77.3%; Score 1350.5; DB 1; Length 303;
 Best Local Similarity 79.7%; Pred. No. 2.5e-76;
 Matches 259; Conservative 9; Mismatches 34; Indels 23; Gaps 3;

QY 1 MSSEAEETQPPAAP-PAAPALSAADTKPGTTGSGAGSGPGGLTSAAPAGGDKKVIATKV 60
 DB 1 MSSEAEETQ-----QQDALEKAGQ-----EPAATVGDKKVIATKV 38
 QY 61 GTVKWFNVNGYGFINRNDTKEDVFVHQTAKKNNPKYLSRSGDGETVEFDVGEKG 120
 DB 39 GTVKWFNVNGYGFINRNDTKEDVFVHQTAKKNNPKYLSRSGDGETVEFDVGEKG 98
 QY 121 EAAANVTGPGVPVQGSKYAADRNHYRYPERRGPPPNYQNNYQNSGKESGSAPE 180
 DB 99 EAAANVTGPGVPVQGSKYAADRNHYRYPERRGPPPNYQNNYQNSGKESGSAPE 158
 QY 181 -QAQRRPYRRRRPPPYMRRPYGRRPYQSNPPVQGEVMEGADNQGAGEQGRVPRQNM 239
 DB 159 DDSNQRRPYRRRRPPPYMRRPYGRRPYQSNPPVQGEVMEGADNQGAGEQGRVPRQNM 218
 QY 240 GYRFRFRGPPRQPRQREDGNEEDKENQDGTQGGPPQRRYRNFYRRRRRPNKPD 299
 DB 219 GYRFRFRGPPRQPRQREDGNEEDKENQDGTQGGPPQRRYRNFYRRRRRPNKPD 278
 QY 300 GKETAADPPAENSSAPEAQGGAE 324
 DB 279 GKETAETSAETSTPEAQGGAE 303

RESULT 5
 ID_YB3_XENLA STANDARD; PRT; 305 AA.
 AC Q00436;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-DEC-1992 (Rel. 24, Last annotation update)
 DE B box binding protein (YB3 protein).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RX MEDLINE=91367675; PubMed=1891365;
 RA Cohen I., Reynolds W.F.;
 RT "The Xenopus YB3 protein binds the B box element of the class III
 promoter.";
 RL Nucleic Acids Res. 19:4753-4759(1991).
 CC -!- FUNCTION: Binds the B box promoter element of genes transcribed
 by RNA polymerase III and is probably involved in the regulation
 of class III genes.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: Belongs to the cold-shock domain (CSD) family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; X60217; CAA42778.1; --
DR PIR; S22313; S22313.
DR HSSP; P41016; IC90.
DR TRANSFAC; T02497; --
DR InterPro; IPR002059; Cold shock.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR Pfam; PF00313; CSD; 1.
DR PRINTS; PR00050; COLDSHOCK.
DR ProDom; PD000621; Cold shock; 1.
DR SMART; SM00357; CSP; 1.
DR PROSITE; PS00352; COLD SHOCK; 1.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT DOMAIN 39 103 CSD.
SQ SEQUENCE 305 AA; 34484 MW; F47F23510FCBBC7F CRC64;

Query Match 74.7%; Score 1304.5; DB 1; Length 305;
Best Local Similarity 78.4%; Pred. No. 1.7e-73;
Matches 258; Conservative 13; Mismatches 29; Indels 29; Gaps 6;

QY 1 MSSEAE--COPPAAPPAALSAADTKPGTTGSGAGSGPGGLTSAAPAGGDKVIATK 58
DB 1 MSSEVEEQEQPDLE-----GKAGEPAATV-GEKKVIATK 36
QY 59 VLGVTKWFNVRNGYGFINRNDTKEDVFVHQTAKKNNPKYLSVSGDGVVEGEK 118
DB 37 VLGVTKWFNVRNGYGFINRNDTKEDVFVHQTAKKNNPKYLSVSGDGVVEGEK 96
QY 119 GAEEAANTGPGGVQVQSKVAADRNHRYPRRRGPPRNVCQNYQNSGEKNGESAP 178
DB 97 GAEEAANTGPGGVQVQSKVAADRNHRYPRRRGPPRNVCQNYQNSGEKNGESAP 156
QY 179 EGO-AQORRRPYRRRF-PPYVMRPPYGRRRPOYSNPPVQGVMEGADNQAGEQRPVRQ 236
DB 157 EGETNQCPCTPRRRYPFVSRYPGRRPQYSNVPVQGESAEAGSQAGEQRPVRQ 216
QY 237 MYRGVRRFRGPRQRPREDGNEEDKENGDEDTQCGQ-PPQRRYRNRYRRRRRPNP 295
DB 217 MYRGVRRFRGPRQRPREDGNEEDKENGDEDTQSHACHLMRYRRNRYRRRRRPNP 276
QY 296 KPQDGKETKAADPPAENSAPAEQGGAE 324
DB 277 KPQDGKETKAETSAPAEQGGAE 305

RESULT 6

ID DBPA RAT STANDARD; PRT; 361 AA.
AC Q62764; Q63748;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA-binding protein A (Cold shock domain protein A) (Muscle Y-box
DE protein YB2) (Y-box binding protein-A) (RYB-A).
GN CSDA OR DBPA OR YB2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley;
RA Goldman D., Gao J., Burneister M., Sapru M.;
RT "Characterization of muscle Y-box proteins that bind the nACHR
RT delta subunit promoter";
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

RC STRAIN-Wistar; TISSUE=Liver;
RX MEDLINE=94301785; PubMed=8029009;
RA Ito K., Tsutsumi K., Kuzumaki T., Gomez P.F., Otsu K., Ishikawa K.;
RT "A novel growth-inducible gene that encodes a protein with a
RT conserved cold-shock domain";
RL Nucleic Acids Res. 22:2036-2041(1994).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=1;
CC IsoId=Q62764-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q62764-2; Sequence=VSP 001137;
CC -1- TISSUE SPECIFICITY: Abundant in the skeletal muscle, spleen and
CC fetal liver.
CC -1- SIMILARITY: Belongs to the cold-shock domain (CSD) family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; U22893; AAB60520.1; --
DR EMBL; D28557; BA05907.1; --
DR HSSP; P41016; IC90.
DR InterPro; IPR002059; Cold shock.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR Pfam; PF00313; CSD; 1.
DR PRINTS; PR00050; COLDSHOCK.
DR ProDom; PD000621; Cold shock; 1.
DR SMART; SM00357; CSP; 1.
DR PROSITE; PS00352; COLD SHOCK; 1.
KW Transcription regulation; Repressor; DNA-binding; Nuclear protein;
KW Alternative splicing.
FT DOMAIN 85 149 CSD.
FT VARSP LIC 184 252 Missing (in isoform 2).
FT FTID=VSP 001137.
FT L->H (IN REF. 2).
FT SPGDAAGPAPASAPAGSEDA -> APARASPRARPOLI
FT SPRGRGR (IN REF. 2).
FT PR -> HV (IN REF. 2).
FT CONFLICT 120 121
SQ SEQUENCE 361 AA; 38851 MW; C6799D5A3DA5C3F3 CRC64;

Query Match 49.5%; Score 865.5; DB 1; Length 361;
Best Local Similarity 55.5%; Pred. No. 1.6e-46;
Matches 193; Conservative 25; Mismatches 101; Indels 29; Gaps 10;

QY 2 SSEAETQPPAAPPAALSAADTKPG-----TTGSGAGSGGPG-GLTSAAPAG---GDKK 53
DB 18 AADAPAAAPPDPAPKSPASGAPAPAPAAALLAGSPGDAAPGAPASAPAGSEDAEKK 77
QY 54 VIATKVLGVTKWFNVRNGYGFINRNDTKEDVFVHQTAKKNNPKYLSVSGDGVVEFDV 113
DB 78 VIATKVLGVTKWFNVRNGYGFINRNDTKEDVFVHQTAKKNNPKYLSVSGDGVVEFDV 137
QY 114 VEGEKGAEEAANTGPGGVQVQSKVAADRNHRYR--YPRRRGPPRNVCQNYQNSGEK 171
DB 138 VEGEKGAEEAANTGPGGVQVQSKVAADRNHRYR--YPRRRGPPRNVCQNYQNSGEK 197
QY 172 EGESAPAGQ-----AQRRPYRRRRPPYMRPYGRPO---YSNPPVQGVEM 218
DB 198 GFEPFAADGQFGARNQLRRPQYRPPYRQRRFPYHVGTGTFDRSRVFPHPNRMQAGEIG 257
QY 219 EGADNQGAGEQGRPVFQNNMYRGVRRFRGPRQRPREDGNEEDKENGDEDTQCGQPPQ 278
DB 258 EMKDGVEGAQ---LQVHRNRYRRFRGPRPAPPAIGAEADKENGQAAANGPNQPSA 314
QY 279 RR-YRRNRYRRR-REENPKPDGKETKAADPPAENSAPAEQGGAE 324


```
QY 44 SAAPAG-----GDKVIATKVLGTWKFNVRNGYGFINRNDTKEDV 86
Db 59 DAAPATGTAASALAAAGSEAEKVKLATKVLGTWKFNVRNGYGFINRNDTKEDV 118
QY 87 HQTAIKNNPRKYLRSVGDGETVEFVVEGEKGAEEANVTGPGVFPVQGSKYAADRNHYR 146
Db 119 HQTAIKNNPRKYLRSVGDGETVEFVVEGEKGAEEANVTGPGVFPVQGSKYAADRNHYR 178
QY 147 R--YPRRGPRPNYQYQNSGEKNEGESAPG-----CAQORRP----- 187
Db 179 RGYGRGRGPRNY-----AGEEEEGSGSEGDPPATDRQFSGARNQLRRFYRP 230
QY 188 -YRRRRPPYMRPFGRRPQ---YSNPPVQGVMEGADN--QGAGEQGRPVQNNMYRGY 241
Db 231 QYQRPRPPYHVGTFDRSRVLPHNPIQAGEIGEMKDGVPESQALQG-PVHNN--PTY 287
QY 242 RPRFR-RGPRQRPRDNEEDKENQDGTGQGPQRR-YRNNFYRRR-RPENPKPQ 298
Db 288 RPYRSRGPRRPAPAVGEAEDKENQATSGNQPVSRRGYRRPYRRRPPPPNAPSQ 347
QY 299 DGKETKAADPPAENSAPAEQGGAE 324
Db 348 DGKEAKAGEAPTEN-PAPPTQSSAE 372

RESULT 8
ID_YB56_XENLA STANDARD; PRT; 336 AA.
AC P21574;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE Cytoplasmic RNA-binding protein p56 (Y box binding protein-2) (Y-box
DE transcription factor) (WRNP4).
GN FRGy2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
EN [1]
EN SEQUENCE FROM N.A.
EN MEDLINE=91062415; PubMed=2247479;
RA Tafari S.R., Wolfe A.P.;
RA "Xenopus Y-box transcription factors: molecular cloning, functional
RA analysis and developmental regulation.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:9028-9032(1990).
RN [2]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP MEDLINE=92107999; PubMed=1729676;
RA Murray M.T., Schiller D.L., Franke W.W.;
RA "Sequence analysis of cytoplasmic mRNA-binding proteins of Xenopus
RA oocytes identifies a family of RNA-binding proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:11-15(1992).
RN [3]
RN PARTIAL SEQUENCE.
RC TISSUE=Ovary;
RC MEDLINE=92332467; PubMed=1629179;
RA Deschamps S., Viel A., Garrigos M., Denis H., le Maire M.;
RA "MRNP4, a major mRNA-binding protein from Xenopus oocytes is
RA identical to transcription factor FRG Y2.";
RL J. Biol. Chem. 267:13799-13802(1992).
RN [4]
RN PARTIAL SEQUENCE.
RC TISSUE=Ovary;
RC MEDLINE=91224309; PubMed=1902800;
RA Deschamps S., Viel A., Denis H., le Maire M.;
RA "Purification of two thermostable components of messenger
RA ribonucleoprotein particles (mRNPs) from Xenopus laevis oocytes,
RA belonging to a novel class of RNA-binding proteins.";
RL FEBS Lett. 282:110-114(1991).
CC -!- FUNCTION: Binds to CCAAT-containing Y box of the hsp70 genes.
CC Seems to be a negative regulatory factor. Also binds to mRNA.
```

```
CC -!- SUBUNIT: Possibly forms a heterodimer with p54 in the 6S and 15S
CC mRNA-binding particles.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic, either free or associated with
CC ribonucleoprotein particles.
CC -!- TISSUE SPECIFICITY: Testis and immature oocytes.
CC -!- PTM: Phosphorylation activates in vitro RNA-binding.
CC -!- SIMILARITY: Belongs to the cold-shock domain (CSD) family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M59454; AA49716.1; -.
CC F01; B38274; B38274.
CC DR HSSP; P41016; IC90.
CC DR TRANSFAC; T00294; -.
CC DR InterPro; IPR002059; Cold shock.
CC DR Pfam; PF00313; CSD; 1.
CC DR PRINTS; PRO0050; COLDSHOCK.
CC DR ProDom; PD00621; Cold_shock; 1.
CC DR SMART; SM00357; CSP; 1.
CC DR PROSITE; PS00352; COLD_SHOCK; 1.
CC KW Transcription regulation; DNA-binding; Nuclear protein;
CC RN-binding; Phosphorylation.
CC FT DOMAIN 44 108
CC FT CONFLICT 254 254 A -> T (IN REF. 1).
CC SQ SEQUENCE 336 AA; 37202 MW; 4AD5838769C6B84D CRC64;
CC
CC Query Match 36.2%; Score 632; DB 1; Length 336;
CC Best Local Similarity 41.8%; Pred. No. 3.1e-32;
CC Matches 152; Conservative 43; Mismatches 91; Indels 78; Gaps 11;
QY 3 SAEATQPPAPPAAPALSAADTKGTGSGAGSGPGLTSAAPAGDKKVIATKVLGT 62
Db 2 SEAAQEPFVP--QPESPEIQLKPG-----IAAARNQANKVLTATVQGT 45
QY 63 VKWNVNNGYGFINRNDTKEDVFVHOTAIKNNPKYLSVGDGETVEFVVEGEKGA 122
Db 46 VKWNVNNGYGFINRNDTKEDVFVHOTAIKNNPKYLSVGDGETVEFVVEGEKGA 105
QY 123 ANVTGPGVFPVQGSKYAADRNHYR---YPRRGPRPNYQYQNSGEKNEGESAP 179
Db 106 ANVTGPGVFPVQGSKYAADRNHYR---YPRRGPRPNYQYQNSGEKNEGESAP 165
QY 180 GQAQORRPYRRRRFPYMRPYGRRPQYSNPPVQGV-EVMEGADNQ----- 228
Db 166 QRPQRRRP-----PPFYRRFRGRGPRPNQQAETVQSENKDPVAPTSEALASGD 219
QY 229 Q-----GRPVQNNMYRGYRPRFRGPPRQRPREDGNEEDKENQDGTGQGP 283
Db 220 DFQRPPPRFRFRFRFRFR-----PAPQQTPEGGDGEAKAESGEDPRPE--PQRQRNR 272
QY 284 NFNYRRR-----PENPKPDGK--ETKAADPPAENS 313
Db 273 PYVQRRRQGTQVAATAQGEKGAFTQHPASEGTSPDPTDDGAPVQSSAPPDGIADT 332
QY 314 SAPE 317
Db 333 PAPE 336

RESULT 9
YB54_XENLA
ID_YB54_XENLA STANDARD; PRT; 324 AA.
AC P4541;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
```



```

Query Match          9.6%; Score 167; DB 1; Length 68;
Best Local Similarity 46.4%; Pred No. 0.0002;
Matches 32; Conservative 14; Mismatches 19; Indels 4; Gaps 1
      58 KVLGTVKVNVNVRNGVFINRNDTKEDVDFVQHTAIKKNPKRVLSRSGVGTEVPDVRGE 117
      2 KTKGVQWFNESKGFGITPADGSKDVFVHFSAIONG-----LKTLAEQNQVFEIQGQ 57
      118 KGAEAAANVT 126
      58 KGPAAAVNVT 66

```

Search completed: August 25, 2004, 00:24:15
Job time : 27 secs

This Page Blank (uspo)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 00:17:20 ; Search time 116 Seconds
(without alignments)
881.275 Million cell updates/sec

Title: US-10-028-415-40
Perfect score: 1747
Sequence: 1 MSSEAEQQPPAAPAL.....AADPPAENSAPAEQGA 324

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : 'SPTEMBL 25':
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phase:
10: sp_plant:
11: sp_podent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rvirus:
16: sp_bacteriap:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	1728	98.9	Q28618 oryctolagus
2	1701	97.4	Q810K5 mus musculus
3	1690	96.7	Q60951 mus musculus
4	1688	96.6	P37752 rattus sp.
5	1626	93.1	Q60950 mus musculus
6	1569	89.8	Q19156 gallus gall
7	1477.5	84.6	Q90376 columba liv
8	1367	78.2	Q8AVK9 xenopus lae
9	1350.5	77.3	Q8AVY9 xenopus lae
10	1321	75.6	Q15325 homo sapien
11	1246	71.3	Q08243 gallus gall
12	1133.5	64.9	Q7ZU03 brachydanio
13	1133	64.9	Q35584 brachydanio
14	1110.5	63.6	O13014 carassius a
15	996	57.0	Q8AXS2 oryzias lat
16	876	50.1	Q9N1Q2 canis famil

17	872.5	49.9	361	11	Q80WG4	Q80WG4 mus musculu
18	871.5	49.9	361	11	Q9EQF8	Q9eqf8 mus musculu
19	865.5	49.5	361	11	Q9JKS3	Q9jks3 mus musculu
20	850.5	48.7	361	11	Q8BP60	Q8bp60 mus musculu
21	807.5	46.2	263	13	Q90WH2	Q90wh2 fugu rubrip
22	792	45.3	304	6	Q9N1Q3	Q9nlq3 canis famil
23	781	44.7	292	11	Q9EQF7	Q9eqf7 mus musculu
24	753.5	43.1	303	4	Q96B76	Q96b76 homo sapien
25	751	43.0	278	13	Q42232	Q42232 columba liv
26	750.5	43.0	303	4	Q96GD7	Q96gd7 homo sapien
27	713	40.8	298	13	Q90650	Q90650 gallus gall
28	708.5	40.6	308	11	Q61478	Q61478 mus musculu
29	687	39.3	274	11	Q64369	Q64369 mus musculu
30	675.5	38.7	360	11	Q922C8	Q922c8 mus musculu
31	637.5	36.5	364	4	Q8N4P0	Q8n4p0 homo sapien
32	637.5	36.1	364	4	Q9Y2T7	Q9y2t7 homo sapien
33	630.5	36.1	282	11	Q922C7	Q922c7 mus musculu
34	620	35.5	199	11	Q9CTJ4	Q9ctt4 mus musculu
35	580.5	33.2	297	13	O13015	O13015 carassius a
36	524.5	30.0	352	5	Q95RE4	Q95re4 drosophila
37	512.5	29.3	317	5	O8IT93	O8it93 chironomus
38	495	28.3	264	5	O8IT94	O8it94 chironomus
39	494.5	28.3	342	5	O61580	O61580 drosophila
40	478	27.4	280	5	O86FT3	O86ft3 dugesia er
41	474.5	27.2	359	5	O46173	O46173 drosophila
42	458.5	26.2	366	13	Q90WH1	Q90wh1 oryzias lat
43	455	26.0	83	11	Q8BRJ5	Q8brj5 mus musculu
44	447	25.6	96	6	Q9XSU1	Q9xsu1 canis famil
45	424.5	24.3	262	13	Q90676	Q90676 gallus gall

ALIGNMENTS

RESULT 1
Q28618 PRELIMINARY; PRT; 324 AA.
ID AC Q28618; PRELIMINARY; PRT; 324 AA.
AC Q28618; PRELIMINARY; PRT; 324 AA.
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE P50.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
ON NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=95155408; PubMed=7852402;
RA Evdokimova V.M., Wei C.L., Sitkov A.S., Simonenko P.N., Lazarev O.A.,
RA Vasilenko K.S., Ustinov V.A., Hershey J.W., Ovchinnikov L.P.;
RT "The major protein of messenger ribonucleoprotein particles in somatic
RT cells is a member of the Y-box binding transcription factor family.";
RL J. Biol. Chem. 270:3186-3192(1995).
CC -!- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
DR EMBL; UI6821; AAA6069.1; -;
DR PIR; A55971; A55971.
DR HSP; P41016; IC90.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR002059; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR008994; Cold shock.
DR Fram; PF00313; CSD; 1.
DR PRINTS; PR00050; COLD SHOCK.
DR PRODOM; PD000621; Cold shock; 1.
DR SMART; SMC0357; CSP; 1.
DR PROSITE; PS00352; COLD SHOCK; 1.
DR SEQUENCE 324 AA; 35824 MW; B2700FD2E61BF8B9 CRC64;

Query Match 98.9%; Score 1728; DB 6; Length 324;
Best local Similarity 98.8%; Pred. No. 1.8e-122; Indels 0; Gaps 0;
Matches 320; Conservative 2; Mismatches 2;

```

QY 1 MSSEAEETQPPAAPALSAADTKPGTTGSGAGSGGGLTSAAPAGDGKKVIATKVL 60
|
|
|
Db 1 MSSEAEETQPPAAPALSAADTKPGTTGSGAGSGGGLTSAAPAGDGKKVIATKVL 60
|
|
|
QY 61 GTVKWNVNRYNGYGFNRNDTKEDVHVHQTAKKNNPKYLRVSGDGETVEFDVVEGEKGA 120
|
|
|
Db 61 GTVKWNVNRYNGYGFNRNDTKEDVHVHQTAKKNNPKYLRVSGDGETVEFDVVEGEKGA 120
|
|
|
QY 121 EAAVNTGPGGVVQGSKYAADRNHYRYPYRRRPPRYNQYQNSSEKNEGSSESAPEG 180
|
|
|
Db 121 EAAVNTGPGGVVQGSKYAADRNHYRYPYRRRPPRYNQYQNSSEKNEGSSESAPEG 180
|
|
|
QY 181 QAQRRPYRRRPPYVYRRPYGRPPQYSNPPVQGEVMEGADNQGAGEQGRPVQRNNMYRG 240
|
|
|
Db 181 QAQRRPYRRRPPYVYRRPYGRPPQYSNPPVQGEVMEGADNQGAGEQGRPVQRNNMYRG 240
|
|
|
QY 241 YRPRFRGPPRQRPREDGNEEDKENOGDETOGQOPQRRYRRNFRNRRRPNPKPDQG 300
|
|
|
Db 241 YRPRFRGPPRQRPREDGNEEDKENOGDETOGQOPQRRYRRNFRNRRRPNPKPDQG 300
|
|
|
QY 301 KETKAADPPAENSAPAEAOGGAE 324
|
|
|
Db 301 KETKAADPPAENSAPAEAOGGAE 324
|
|
|

RESULT 2
ID Q810K5 PRELIMINARY; PRT; 322 AA.
AC Q810K5;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC049977; AAH49977.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR Pfam; PF00313; CSD; 1.
DR PRINTS; PR00050; COLDSHOCK.
DR ProDom; PD000621; Cold_shock; 1.
DR SMART; SM00357; CSP; 1.
DR PROSITE; PS00352; COLD_SHOCK; 1.
KW Hypothetical protein.
SQ SEQUENCE 322 AA; 35756 MW; D7114692FC51A654 CRC64;

Query Match 97.4%; Score 1701; DB 11; Length 322;
Best Local Similarity 98.1%; Pred. No. 2e-120;
Matches 318; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 1 MSSEAEETQPPAAPALSAADTKPGTTGSGAGSGGGLTSAAPAGDGKKVIATKVL 60
|
|
|
Db 1 MSSEAEETQPPAAPALSAADTKPGTTGSGAGSGGGLTSAAPAGDGKKVIATKVL 58
|
|
|
QY 61 GTVKWNVNRYNGYGFNRNDTKEDVHVHQTAKKNNPKYLRVSGDGETVEFDVVEGEKGA 120
|
|
|
Db 59 GTVKWNVNRYNGYGFNRNDTKEDVHVHQTAKKNNPKYLRVSGDGETVEFDVVEGEKGA 118
|
|
|
QY 121 EAAVNTGPGGVVQGSKYAADRNHYRYPYRRRPPRYNQYQNSSEKNEGSSESAPEG 180
|
|
|
Db 121 EAAVNTGPGGVVQGSKYAADRNHYRYPYRRRPPRYNQYQNSSEKNEGSSESAPEG 180
|
|
|
QY 181 QAQRRPYRRRPPYVYRRPYGRPPQYSNPPVQGEVMEGADNQGAGEQGRPVQRNNMYRG 240
|
|
|
Db 181 QAQRRPYRRRPPYVYRRPYGRPPQYSNPPVQGEVMEGADNQGAGEQGRPVQRNNMYRG 240
|
|
|
QY 241 YRPRFRGPPRQRPREDGNEEDKENOGDETOGQOPQRRYRRNFRNRRRPNPKPDQG 300
|
|
|
Db 241 YRPRFRGPPRQRPREDGNEEDKENOGDETOGQOPQRRYRRNFRNRRRPNPKPDQG 300
|
|
|
QY 301 KETKAADPPAENSAPAEAOGGAE 324
|
|
|
Db 301 KETKAADPPAENSAPAEAOGGAE 324
|
|
|

RESULT 4
ID Q60951 PRELIMINARY; PRT; 322 AA.
AC Q60951;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE MYB-1b.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57 Black/6 X CBA; TISSUE=Thymus;
RA Schwartz B.D.; Meador J.W. III; Woulfe S.L.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
DR EMBL; U33197; AA875476.1; -.
DR HSSP; F41016; IC90.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR002059; Cold_shock.
DR Pfam; PF00313; CSD; 1.
DR PRINTS; PR00050; COLDSHOCK.
DR ProDom; PD000621; Cold_shock; 1.
DR SMART; SM00357; CSP; 1.
DR PROSITE; PS00352; COLD_SHOCK; 1.
SQ SEQUENCE 322 AA; 35700 MW; 3F216C42CA1671BB CRC64;

Query Match 96.7%; Score 1690; DB 11; Length 322;
Best Local Similarity 97.8%; Pred. No. 1.3e-119;
Matches 317; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

QY 1 MSSEAEETQPPAAPALSAADTKPGTTGSGAGSGGGLTSAAPAGDGKKVIATKVL 60
|
|
|
Db 1 MSSEAEETQPPAAPALSAADTKPGTTGSGAGSGGGLTSAAPAGDGKKVIATKVL 58
|
|
|
QY 61 GTVKWNVNRYNGYGFNRNDTKEDVHVHQTAKKNNPKYLRVSGDGETVEFDVVEGEKGA 120
|
|
|
Db 59 GTVKWNVNRYNGYGFNRNDTKEDVHVHQTAKKNNPKYLRVSGDGETVEFDVVEGEKGA 118
|
|
|
QY 121 EAAVNTGPGGVVQGSKYAADRNHYRYPYRRRPPRYNQYQNSSEKNEGSSESAPEG 180
|
|
|
Db 119 EAAVNTGPGGVVQGSKYAADRNHYRYPYRRRPPRYNQYQNSSEKNEGSSESAPEG 178
|
|
|
QY 181 QAQRRPYRRRPPYVYRRPYGRPPQYSNPPVQGEVMEGADNQGAGEQGRPVQRNNMYRG 240
|
|
|
Db 179 QAQRRPYRRRPPYVYRRPYGRPPQYSNPPVQGEVMEGADNQGAGEQGRPVQRNNMYRG 238
|
|
|
QY 241 YRPRFRGPPRQRPREDGNEEDKENOGDETOGQOPQRRYRRNFRNRRRPNPKPDQG 300
|
|
|
Db 239 YRPRFRGPPRQRPREDGNEEDKENOGDETOGQOPQRRYRRNFRNRRRPNPKPDQG 298
|
|
|
QY 301 KETKAADPPAENSAPAEAOGGAE 324
|
|
|
Db 299 KETKAADPPAENSAPAEAOGGAE 322
|
|
|

RESULT 4

```

P97752 PRELIMINARY; PRT; 322 AA.
AC P97752;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE TSH receptor suppressor element-binding protein-1.
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96435261; PubMed=8838147;
RA Omori M., Shimura H., Shimura Y., Kohn L.D.;
RT "A Y-box protein is a suppressor factor that decreases thyrotropin
RT receptor gene expression.";
RL Mol. Endocrinol. 10:78-89 (1996).
CC -!- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
DR EMBL; S83025; AAB46889.2; -;
DR HSP; P41016; 1C90.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR008994; Cold shock.
DR Pfam; PF00313; CSD; 1.
DR PRINTS; PR00050; COLDSHOCK.
DR ProDom; PD000621; Cold shock; 1.
DR SMART; SM00357; CSP; 1-
DR PROSITE; PS00352; COLD_SHOCK; 1.
KW Receptor.
SQ SEQUENCE 322 AA; 35531 MW; DD6682F62F11D4A8 CRC64;

Query Match 96.6%; Score 1688; DB 11; Length 322;
Best Local Similarity 97.5%; Pred. No. 1.9e-119;
Matches 316; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

Qy 1 MSSAEATQPPAAPAAPALSAADTKPGTTGSGAGSGGPGGLTSAAPAGGDKVIATKVL 60
Db 1 MSSAEATQPPAAP--AAALSAADTKPGSTGSGAGSGGPGGLTSAAPAGGDKVIATKVL 58

Qy 61 GTVKWNVNRNGYGFINRNDTKEDVFVHQTAKKNPKYLSVGDGTFVDFVVEGEKGA 120
Db 59 GTVKWNVNRNGYGFINRNDTKEDVFVHQTAKKNPKYLSVGDGTFVDFVVEGEKGA 118

Qy 121 EAAVNTGGVGVQGSKYAADRNHYRRYRRRPPRRYQNYQNSGEKNEGSAPEG 180
Db 119 EAAVNTGGVGVQGSKYAADRNHYRRYRRRPPRRYQNYQNSGEKNEGSAPEG 178

Qy 181 QAOQRPRRRRPPYMRPYRRPQYNSNPVQGVMEGADNQGAGEQGRPVRRQNYRG 240
Db 179 QAOQRPRRRRPPYMRPYRRPQYNSNPVQGVMEGADNQGAGEQGRPVRRQNYRG 238

Qy 241 YRPRFRGPPRQRPREDGNEEDKENQDGTQGGQPPQRRYRRNFYRRRRPENPKPDG 300
Db 239 YRPRFRGPPRQRPREDGNEEDKENQDGTQGGQPPQRRYRRNFYRRRRPENPKPDG 298

Qy 301 KETKAADPPAENSAPAEQGGAE 324
Db 299 KETKAADPPAENSAPAEQGGAE 322

RESULT 5
Q60950 PRELIMINARY; PRT; 322 AA.
ID Q60950;
AC Q60950;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MYB-1a.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CS7 Black/6 X CBA; TISSUE=Thymus;
RA Schwartz B.D., Meador J.W. III, Wolfe S.L.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
DR EMBL; U33196; AAA75475.1; -;
DR HSP; P41016; 1C90.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR002059; Cold shock.
DR Pfam; PF00313; CSD; 1.
DR PRINTS; PR00050; COLDSHOCK.
DR ProDom; PD000621; Cold shock; 1.
DR SMART; SM00357; CSP; 1-
DR PROSITE; PS00352; COLD_SHOCK; 1.
SQ SEQUENCE 322 AA; 36007 MW; E4E96AF7924544DA CRC64;

Query Match 93.1%; Score 1626; DB 11; Length 322;
Best Local Similarity 94.4%; Pred. No. 8.9e-115;
Matches 306; Conservative 5; Mismatches 11; Indels 2; Gaps 1;

Qy 1 MSSAEATQPPAAPAAPALSAADTKPGTTGSGAGSGGPGGLTSAAPAGGDKVIATKVL 60
Db 1 MSSAEATQPPAAP--AAALSAADTKPGSTGSGAGSGGPGGLTSAAPAGGDKVIATKVL 58

Qy 61 GTVKWNVNRNGYGFINRNDTKEDVFVHQTAKKNPKYLSVGDGTFVDFVVEGEKGA 120
Db 59 GTVKWNVNRNGYGFINRNDTKEDVFVHQTAKKNPKYLSVGDGTFVDFVVEGEKGA 118

Qy 121 EAAVNTGGVGVQGSKYAADRNHYRRYRRRPPRRYQNYQNSGEKNEGSAPEG 180
Db 119 EAAVNTGGVGVQGSKYAADRNHYRRYRRRPPRRYQNYQNSGEKNEGSAPEG 178

Qy 181 QAOQRPRRRRPPYMRPYRRPQYNSNPVQGVMEGADNQGAGEQGRPVRRQNYRG 240
Db 179 QAOQRPRRRRPPYMRPYRRPQYNSNPVQGVMEGADNQGAGEQGRPVRRQNYRG 238

Qy 241 YRPRFRGPPRQRPREDGNEEDKENQDGTQGGQPPQRRYRRNFYRRRRPENPKPDG 300
Db 239 YRPRFRGPPRQRPREDGNEEDKENQDGTQGGQPPQRRYRRNFYRRRRPENPKPDG 298

Qy 301 KETKAADPPAENSAPAEQGGAE 324
Db 299 KETKAADPPAENSAPAEQGGAE 322

RESULT 6
Q91956 PRELIMINARY; PRT; 326 AA.
ID Q91956;
AC Q91956;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mouse dbpB/VB-1 protein homologue, clone CLFEST34.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Lens fibers;
RA MEDLINE=96437509; PubMed=6840185;
RA Sawada K., Agata K., Eguchi G.;
RT "Characterization of terminally differentiated cell state by
RT categorizing cDNA clones derived from chicken lens fibers.";
RL Int. J. Dev. Biol. 40:531-535 (1996).
CC -!- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
DR EMBL; D26328; BAA05380.1; -;

Query Match		84.6%	Score 1477.5; DB 13; Length 287;
Best Local Similarity		84.6%	Pred. No. 1.2e-103;
Matches 274; Conservative		8; Mismatches 5; Indels 37; Gaps 1;	
QY	1	MSSAEATQOPPAAPAAAPALSAADTKGTGTCGAGSGPGLTSAAPAGDGKKVIATKVL	60
DB	1	MSSEATQ-----PPAGDGKKVIATKVL	23
QY	61	GTVWFWNRVNGYGFINRNDTKEDVFVHQTAIKKNNPKYLSVGDGTVEFDVVVEGEKGA	120
DB	24	GTVWFWNRVNGYGFINRNDTKEDVFVHQTAIKKNNPKYLSVGDGTVEFDVVVEGEKGA	83
QY	121	EAAVNTGGVGVVUGSKYAADRNHYRYPRRRGPPRYQOYNQNSGEKNGESAP	180
DB	84	EAAVNTGGVGVVUGSKYAADRNHYRYPRRRGPPRYQOYNQNSGEKNGEAENIPEG	143
QY	181	QAQORRYRRRRPPPYMYRPFYGRPPSYNPPVQGEVMEGADNOCAGEQGRPVQRNNMYRG	240
DB	144	QOQRRYRYYRRYPYMYRPFYGRPPYQYTWPPVQGEIVGADNOCAGEQGRPVQRNNMYRG	203
QY	241	YRPFRRGPPRQRPREDGNEEDKENOGDETQGOQPPQRRYRNFNRYRRRPNPKPDQG	300
DB	204	YRPFRRGPPRQRPREDGNEEDKENOGDETQSOQPPQRRYRNFNRYRRRPNPKPDQG	263
QY	301	KETKAADPPAENSSAPEAEOGGAE	324
DB	264	KETKAEPPIAENSSAPEAEOGGAE	287
RESULT 8			
Q8AVK9 PRELIMINARY; PRT; 304 AA.			
ID	Q8AVK9		
AC	Q8AVK9		
DT	01-MAR-2003 (TrEMBLrel. 23, Created)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	Similar to nuclease sensitive element binding protein 1.		
OS	Xenopus laevis (African clawed frog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;		
OC	Xenopodinae; Xenopus.		
OX	NCBI_Taxid=8355;		
OX	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Embryo;		
RA	Klein S., Strausberg R.;		
RL	Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC042217; AH42217.1; .		
DR	GO; GO:0003677; F:DNA binding; IEA.		
DR	GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.		
DR	InterPro; IPR002059; Cold shock.		
DR	InterPro; IPR008994; Nucleic_acid_OB.		
DR	Pfam; PF00313; CSD; 1.		
DR	PRINTS; PR00050; COLDSHOCK.		
DR	ProDom; PD000621; Cold_shock; 1.		
DR	SMART; SMC0357; CSP; 1.		
DR	PROSITE; PS00352; COLD_SHOCK; 1.		
SQ	SEQUENCE 304 AA; 34390 MW; 085C185E4813129E CRC64;		
Query Match			
Best Local Similarity		80.8%; Score 1367; DB 13; Length 304;	
Matches 265; Conservative		13; Mismatches 22; Indels 28; Gaps 5;	
QY	1	MSSEAT--QQPPAAAPPAALSAADTKPCTGTCGAGSGPGLTSAAPAGDGKKVIATK	59
DB	1	MSSEVETQQQPPDALE-----CKAQEPAATV-GEKKVIATK	36
QY	59	VLGTWFWNRVNGYGFINRNDTKEDVFVHQTAIKKNNPKYLSVGDGTVEFDVVVEGEK	118
DB	37	VLGTWFWNRVNGYGFINRNDTKEDVFVHQTAIKKNNPKYLSVGDGTVEFDVVVEGEK	96
QY	119	GAEAAVNTGGVGVVUGSKYAADRNHYRYPRRRGPPRYQOYNQNSGEKNGESAP	178

```

Db 97 GAFAANVTGGVVOGSKYAADNNRYRRRRRPPRYQQNYQNSSEKAEAGNESAP 156
QY 179 EGO-AQORRRYRRRF-PPYMRPPYGRPPQYGNPPVQGVMBGADNQGAGQGRVQRN 236
Db 157 EGGTQOQYPPRRYPPYPPYRRPPYGRPPQYGNPPVQGVMBGADNQGAGQGRVQRN 216
QY 237 MYRGYPRFRPPRRPPQRPREDNEEDKENQGDQGGTQGGPPQRRYRNRYRRRRPENPK 296
Db 217 MYRGYPRFRPPRRPPQRPREDNEEDKENQGDQGGTQGGPPQRRYRNRYRRRRPENPK 276
QY 297 PQGKETKAADPPAENSSAPEAEOGGAE 324
Db 277 PQGKETKAADPPAENSSAPEAEOGGAE 304

RESULT 9
Q8AVY9 PRELIMINARY; PRT; 303 AA.
AC Q8AVY9; 2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Similar to nuclease sensitive element binding protein 1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Klein S.; Strausberg R.; to the EMBL/GenBank/DBJ databases.
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041191; AAH41191.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR02059; Cold shock.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR Pfam; PF00313; CSD; 1.
DR PRINTS; PR00050; COLDSHOCK.
DR ProDom; PD000621; Cold shock; 1.
DR SMART; SM00357; CSP; 1.
DR PROSITE; PS00352; COLD SHOCK; 1.
SQ SEQUENCE 303 AA; 34647 MW; 6D0E8AB4257134D4 CRC64;

Query Match 77.3%; Score 1350.5; DB 13; Length 303;
Best Local Similarity 79.7%; Pred. No. 5e-94;
Matches 259; Conservative 9; Mismatches 34; Indels 23; Gaps 3;

QY 1 MSSEATQOPPAAPPAALSAADTKPGTTGSGAGSGGPGGLTSAAPAGDKKVIATKVL 60
Db 1 MSSEVETQ-----QPDALGKAGQ-----EPAATVGDKKVIATKVL 38
QY 61 GTVKWNVNRYGFINRNDTKEDVFVHQTAKKNNPKYLRVSGDGETVEFDVVEGEKA 120
Db 39 GTVKWNVNRYGFINRNDTKEDVFVHQTAKKNNPKYLRVSGDGETVEFDVVEGEKA 98
QY 121 EAANVTGPGVVOGSKYAADNNRYRRRRRPPRYQQNYQNSSEKAEAGNESAP 180
Db 99 EAANVTGPGVVOGSKYAADNNRYRRRRRPPRYQQNYQNSSEKAEAGNESAP 158
QY 181 -QAQORRRYRRRRPPYMRPPYGRPPQYGNPPVQGVMBGADNQGAGQGRVQRNMYR 239
Db 159 DDSNQORRRYRRRRPPYMRPPYGRPPQYGNPPVQGVMBGADNQGAGQGRVQRNMYR 218
QY 240 GYRFRFRPPRRPPQRPREDNEEDKENQGDQGGTQGGPPQRRYRNRYRRRRPENPKPD 299
Db 219 GYRFRFRPPRRPPQRPREDNEEDKENQGDQGGTQGGPPQRRYRNRYRRRRPENPKPD 278
QY 300 GKETKAADPPAENSSAPEAEOGGAE 324
Db 279 GKETKAADPPAENSSAPEAEOGGAE 303

```

```

RESULT 10
Q15325 PRELIMINARY; PRT; 320 AA.
ID Q15325; Q14972;
AC Q15325; Q14972;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Nuclease sensitive element binding protein 1.
GN NSEP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kolluri R.; Torrey T.A.; Davis-Smythe T.L.; Kinniburgh A.J.;
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=92150140; PubMed=1738588;
RA Kolluri R.; Torrey T.A.; Kinniburgh A.J.;
RT "A CT promoter element binding protein: definition of a double-strand
RT and a novel single-strand DNA binding motif.";
RL Nucleic Acids Res. 20:111-116(1992).
CC 1- FUNCTION: BINDS CT-RICH DUPLEX AND CT-RICH SINGLE STRAND DNA.
CC 2- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
DR EMBL; M83234; AAA59949.1; -.
DR HSSP; P41016; 1C90.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0003690; F:double-stranded DNA binding; TAS.
DR GO; GO:0003697; F:single-stranded DNA binding; TAS.
DR GO; GO:0003700; F:transcription factor activity; TAS.
DR GO; GO:0009613; P:response to pest/pathogen/parasite; TAS.
DR GO; GO:0006366; P:transcription from Pol II promoter; TAS.
DR InterPro; IPR02059; Cold shock.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR Pfam; PF00313; CSD; 1.
DR PRINTS; PR00050; COLDSHOCK.
DR ProDom; PD000621; Cold shock; 1.
DR SMART; SM00357; CSP; 1.
DR PROSITE; PS00352; COLD_SHOCK; 1.
KW DNA-binding.
FT DOMAIN 1 57 NECESSARY FOR SINGLE-STRAND DNA-BINDING.
FT DOMAIN 1 75 NECESSARY FOR SINGLE-STRAND AND DOUBLE-
FT DOMAIN 58 122 STRAND DNA-BINDING.
FT CONFLICT 21 47 RPTTSPALRRRRRWPGRLLTSAALRR -> SAADTKPGTT
FT CONFLICT 183 188 GAAQAVARAHAHIGAGAG (IN REF. 2).
FT CONFLICT 320 AA; 35368 MW; F534DB708FDCAE8D CRC64;
SQ SEQUENCE 320 AA; 35368 MW; F534DB708FDCAE8D CRC64;

Query Match 75.6%; Score 1321; DB 4; Length 320;
Best Local Similarity 81.5%; Pred. No. 8.9e-92;
Matches 264; Conservative 2; Mismatches 54; Indels 4; Gaps 3;

QY 1 MSSEATQOPPAAPPAALSAADTKPGTTGSGAGSGGPGGLTSAAPAGDKKVIATKVL 60
Db 1 MSSEATQOPPAAPPAALSAADTKPGTTGSGAGSGGPGGLTSAAPAGDKKVIATKVL 57
QY 61 GTVKWNVNRYGFINRNDTKEDVFVHQTAKKNNPKYLRVSGDGETVEFDVVEGEKA 120
Db 58 GTVKWNVNRYGFINRNDTKEDVFVHQTAKKNNPKYLRVSGDGETVEFDVVEGEKA 117
QY 121 EAANVTGPGVVOGSKYAADNNRYRRRRRPPRYQQNYQNSSEKAEAGNESAP 180

```

```
Db 118 EAAVTGPGVPGVGSKYAADRNHYRRYPRRRGPRNYQQYQNSGEKNESESAPAE 177
Qy 181 QAOQRPRYRRRFPYRRPQYSNPPVOGEVMEGADNOGAGEQGRPVYRQNMRYG 240
Db 178 RPNNAAY-AGEVTLTHAETLGRPOYSNPPVOGEVMEGADNOGAGEQGRPVYRQICIGD 236
Qy 241 YRPRFRGPPRRQPRREDGNEEDKENQDGTQGGQPPQRRYRNFYRRRPNPKPDG 300
Db 237 IDHDSAGALLAKROFREDGNEEDKENQDGTQGGQPPQRRYRNFYRRRPNPKPDG 296
Qy 301 KETKAADPPAENSSAPAEQGGAE 324
Db 297 KETKAADPPAENSSAPAEQGGAE 320

RESULT 11
Q08243 PRELIMINARY; PRT; 348 AA.
AC Q08243,
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE CCAAT-binding protein (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94120725; PubMed=8291233;
RA Kandala J.C., Guntaka R.V.;
RT "Cloning of Rous sarcoma virus enhancer factor genes. I. Evidence that
RT RSV-EF-I is related to Y-box (inverted CCAAT) binding proteins and
RT binds to multiple motifs in the RSV enhancer."
RL Virology 198;1514-523(1994).
CC -!- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
DR EMBL; L20500; AAA18017.1; -
DR PIR; A49594; A49594.
DR HSP; P41016; LC90.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004601; F:peroxidase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000028; Chloroperoxidase.
DR InterPro; IPR002059; Cold shock.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR Pfam; PF00313; CSD; 1.
DR ProDom; PD000621; Cold shock; 1.
DR SMART; SMC0357; CSP; 1.
FT NON TER 1
SQ SEQUENCE 348 AA; 39405 MW; 4795EA2007ED062F CRC64;

Query Match 71.3%; Score 1246; DB 13; Length 348;
Best Local Similarity 75.0%; Pred. No. 4.5e-86;
Matches 243; Conservative 15; Mismatches 48; Indels 18; Gaps 4;

Qy 5 AETQPP-----PAAPPAAPALSAADTKPGTTGSGAGSGGPGGLTSAAPAGDKKVIATK 58
Db 4 AETQPPAAPVPAPPPPPPTTNTLAA-----ACTGAAMPRLRLRERGQEGH-----RNE 52
Qy 59 VLGTWKFVNRVNGYGFINRNDTKEDVFVHQTAKKNNPKYLRVSGDGTVEFDVVEGK 118
Db 53 GFGNSEMVNRVNGYGLMNRNDTKEDVFVHQTAKKNNPKYLRVSGDGTVEFDVVEGK 112
Qy 119 GAAANVTGPGVPGVGSKYAADRNHYRRYPRRRGPRNYQQYQNSGEKNESESAP 178
Db 113 GAAANVTGPGVPGVGGKVAADRNHYRRYPRVAGVLHATSKTENSESGEKNEGAENIP 172
Qy 179 EGQAQRRPYRRRFPYRRPQYSNPPVOGEVMEGADNOGAGEQGRPVYRQNMRYG 238
Db 173 EGQAQRRPYRRRFPYRRPQYSNPPVOGEVMEGADNOGAGEQGRPVYRQNMRYG 231
```

```
Qy 239 RGYRPRFRGPPRRQPRREDGNEEDKENQDGTQGGQPPQRRYRNFYRRRPNPKPQ 298
Db 232 RGYRPRFRGPPRRQPRREDGNEEDKENQDGTQGGQPPQRRYRNFYRRRPNPKPQ 291
Qy 299 DKETKAADPPAENSSAPAEQGG 322
Db 292 DKETKAADPPAENSSAPAEQAG 315

RESULT 12
Q7ZU03 PRELIMINARY; PRT; 309 AA.
AC Q7ZU03,
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DE Nuclease sensitive element binding protein 1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Body;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Locquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.B.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Body;
RA Strausberg R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC050156; AAH50156.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR002059; Cold shock.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR Pfam; PF00313; CSD; 1.
DR PRINTS; PR00050; COLDSHOCK.
DR ProDom; PD000621; Cold shock; 1.
DR SMART; SM00357; CSP; 1.
DR PROSITE; PS00352; COLD_SHOCK; 1.
SQ SEQUENCE 309 AA; 34935 MW; 7CF1CCFE7DF5777A CRC64;

Query Match 64.9%; Score 1133.5; DB 13; Length 309;
Best Local Similarity 69.6%; Pred. No. 1.2e-77;
Matches 233; Conservative 26; Mismatches 39; Indels 37; Gaps 11;

Qy 1 MSSAEATQPPAPPAAPALSAADTKPGTTGSGAGSGGPGGLTSAAPAGDKKVIATKVL 60
Db 1 MSSAEATQPPAPPAAPALSAADTKPGTTGSGAGSGGPGGLTSAAPAGDKKVIATKVL 40
Qy 61 GTVWFVNRVNGYGFINRNDTKEDVFVHQTAKKNNPKYLRVSGDGTVEFDVVEGKGA 120
```



```
Db 41 GTVKWNVNRYGFIINRNDTKEDVFHQTAIKNNPRKYLRSVGDGETVEFDVVEGEKA 100
QY 121 EAAVNTGPGGVPVQGSKYAADRNHYRRYPRRRGPPRNYQNYQN---SESGKNESGESA 177
Db 101 EAAVNTGPGGVPVQGSKYAADRNHYRRYPRRRGPPRNYQNYQSDPEAPREKREGAESA 160
QY 178 PEG--QAQORRPY--RRRRPPPYMRPPYGRRRPOYSNPPVQGVMEGA---DNQAGEQG 230
Db 161 PEGEMOQQQRPTYPGRRRYPPYFVRRYGRPPYTNQ-RGEMTGGGSEENQGGPDQG 219
QY 231 -RPVQNMYRGYRPRRRGPPRQRPREDGNEEDKENQGDDETOGQPPQRRYRNFNRYR 289
Db 220 NKPMQNYRGFRP---RGPPRR--PVRDG-EEDKENQSGQNPQRRYRNFNRYR 274
QY 290 RRPENPKQCKTKKAADPPAENSAPAEBOGGAE 324
Db 275 RRPQTTFQDQKSKAADASADKSAAPAEBOGGAD 309
RESULT 13
O93584 PRELIMINARY; PRT; 310 AA.
AC O93584;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Zfy1.
GN nsepl.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99376608; PubMed=10446383;
RA Chang B.E., Lin C.Y., Kuo C.M.;
RT "Molecular cloning of a cold-shock domain protein, zfy1, in zebrafish
embryo."
RL Biochim. Biophys. Acta 1433:343-349(1999).
CC 1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
DR EMBL; AF093129; AAC62774.1; -.
DR HSP; P41016; LC90.
DR ZFIN; ZDB-GENE-000629-3; nsepl.
DR GO; GO:0003677; P:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR002059; Cold shock
DR InterPro; IPR008994; Nucleic_acid_OB.
DR Pfam; PF00313; CSD; 1.
DR ProDom; PD000621; Cold shock; 1.
DR PRINTS; PR00050; COLDSHOCK.
DR SMART; SM00357; CSP; 1.
DR PROSITE; PS00352; COLD_SHOCK; 1.
SQ SEQUENCE 310 AA; 35022 MW; 64430D39DD56D0DE CRC64;
Query Match 64.9%; Score 1133; DB 13; Length 310;
Best Local Similarity 69.6%; Pred. No. 1.3e-77;
Matches 233; Conservative 26; Mismatches 40; Indels 36; Gaps 11;
QY 1 MSSAEATQPPAAPAALSAADTKPGTTGSGAGSGPGGLTSAAPAGDKKVIATKVL 60
Db 1 MSSAEATQPP--QFADAESFS-----PAAATAGDKKVIATKVL 40
QY 61 GTVKWNVNRYGFIINRNDTKEDVFHQTAIKNNPRKYLRSVGDGETVEFDVVEGEKA 120
Db 41 GTVKWNVNRYGFIINRNDTKEDVFHQTAIKNNPRKYLRSVGDGETVEFDVVEGEKA 100
QY 121 EAAVNTGPGGVPVQGSKYAADRNHYRRYPRRRGPPRNYQNYQN---SESGKNESGESA 177
Db 101 EAAVNTGPGGVPVQGSKYAADRNHYRRYPRRRGPPRNYQNYQSDPEAPREKREGAESA 160
QY 178 PEG--QAQORRPY--RRRRPPPYMRPPYGRRRPOYSNPPVQGVMEGA---DNQAGEQG 230
```

```
Db 161 PEGEMOQQQRPTYPGRRRYPPYFVRRYGRPPYTNQ-RGEMTGGGSEENQGGPDQG 219
QY 231 -RPVQNMYRGYRPRRRGPPRQRPREDGNEEDKENQGDDETOGQPPQRRYRNFNRYR 289
Db 220 NKPMQNYRGFRP---SRGPPRR--PVRDG-EEDKENQSGQNPQRRYRNFNRYR 275
QY 290 RRPENPKQCKTKKAADPPAENSAPAEBOGGAE 324
Db 275 RRPQTTFQDQKSKAADASADKSAAPAEBOGGAD 310
RESULT 14
O13014 PRELIMINARY; PRT; 311 AA.
AC O13014;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Y box protein 1.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RX Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RA katsu Y.;
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98055714; PubMed=9395336;
RA Katsu Y., Yamashita M., Nagahama Y.;
RT "Isolation and characterization of goldfish Y box protein, a germ-
cell-specific RNA-binding protein."
RL Eur. J. Biochem. 249:1854-1861(1997).
CC 1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
DR EMBL; AB003335; BAA19849.1; -.
DR HSP; P41016; LC90.
DR GO; GO:0003677; P:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR002059; Cold shock.
DR Pfam; PF00313; CSD; 1.
DR PRINTS; PR00050; COLDSHOCK.
DR ProDom; PD000621; Cold shock; 1.
DR SMART; SM00357; CSP; 1.
DR PROSITE; PS00352; COLD_SHOCK; 1.
SQ SEQUENCE 311 AA; 35076 MW; 3611B8F4E5914039 CRC64;
Query Match 63.6%; Score 1110.5; DB 13; Length 311;
Best Local Similarity 68.5%; Pred. No. 6.5e-76;
Matches 228; Conservative 24; Mismatches 50; Indels 31; Gaps 8;
QY 1 MSSAEATQPPAAPAALSAADTKPGTTGSGAGSGPGGLTSAAPAGDKKVIATKVL 60
Db 1 MSSAEATQPP--QFADAES-----PSSPSPAATAGDKKVIATKVL 41
QY 61 GTVKWNVNRYGFIINRNDTKEDVFHQTAIKNNPRKYLRSVGDGETVEFDVVEGEKA 120
Db 42 GTVKWNVNRYGFIINRNDTKEDVFHQTAIKNNPRKYLRSVGDGETVEFDVVEGEKA 101
QY 121 EAAVNTGPGGVPVQGSKYAADRNHYRRYPRRRGPPRNYQNYQ---NSESGKNESGESA 179
Db 102 EAAVNTGPGGVPVQGSKYAADRNHYRRYPRRRGPPRNYQNYQSDPEAPREKREGAESA 161
QY 180 G--QAQORRPY--RRRRPPPYMRPPYGRRRPOYSNPPVQGVMEGA---DNQAGEQG-R 231
Db 162 GEMQQRPTYPGRRRYPPYFVRRYGRPPYTNAPQEGEMPEGEGDENQGVDPQGNK 221
QY 232 PVRQNMYRGYRPRRRGPPRQRPREDGNEEDKENQGDDETOGQPPQRRYRNFNRYR 291
```

Db 222 PMRQNYRGGFGP---RGLPRPRVREGEEKEKQGGQNGQEPFCRCYRRNFYRRR 278
Qy 292 PENPKPDQKTKAADPAENSSAPAEAOGGAE 324
Db 279 PÖTKLOQDKSKAADASADKPAAPAEAOGGAE 311

RESULT 15

Q8AXS2

ID Q8AXS2 PRELIMINARY; PRT; 306 AA.

AC Q8AXS2

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Y-box binding protein.

GN YB1

OS Oryzias latipes (Medaka fish) (Japanese ricefish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;

OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.

OX NCBI_TaxID=8090;

RN [1]

RP SEQUENCE FROM N.A.

RA Ohe N.;

RT "Cloning of Japanese Medaka Y-box binding protein cDNA";

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

RL EMBL; AB041646; BAC45236.1; ..

DR GO: 0003677; F:DNA binding; IEA.

DR GO: 0006355; P:regulation of transcription, DNA-dependent; IEA.

DR InterPro; IPR02059; Cold_shock.

DR Pfam; PF00313; CSD; 1.

DR PRINTS; PR00050; COLDSHOCK.

DR PRODOM; PD00621; Cold shock; 1.

DR SMART; SM00357; CSP; 1.

DR PROSITE; PS00352; COLD_SHOCK; 1.

SQ SEQUENCE 306 AA; 33718 MW; 5244E7EC94D41696 CRC64;

Query Match

Best Local Similarity 57.0%; Score 996; DB 13; Length 306;

Matches 219; Conservative 19; Mismatches 47; Indels 60; Gaps 14;

Qy 1 MSSEAETQPPAPPAALSAADTKPGTTGSGAGSGPGGLTSAAPAGDKKVIATKVL 60

Db 1 MSSEAETQPPAPPAALSAADTKPGTTGSGAGSGPGGLTSAAPAGDKKVIATKVL 42

Qy 61 GTVKNFVNRNGYGFINRNDKEDVFVHQTAKNNPKYLSVGDGETVEFDVVEGEKA 120

Db 43 GTVKNFVNRNGYGFINRNDKEDVFVHQTAKNNPKYLSVGDGETVEFDVVEGEKA 102

Qy 121 EAAVTGPGGVPVGGSYAADRNHYRRYPRRRGPPR--NYQNYQNSSEGE----- 169

Db 103 EAAVTGPGGVPVGGSYAADRNHYRRYPRRRGPPR--NYQNYQNSSEGE----- 162

Qy 170 KNEGESAPEG--QAQORRPY--RRRRFPYMRFPYGRFPQYNSFPVQGEVMEGADNOG 225

Db 163 NRDGESAPEGEQPQORRTYPSRRYPYP-----GVEGDENOG 203

Qy 226 AGEGG-RPVQNMYRGYRPRFR-RGPPRQRPREDNEEDKENQGDG-TGQQPPQRRYR 282

Db 204 GPDQGNKPVQNYRGYRPRFRFTPRGPPRPR-PVRDG-EEDKENQGGGQNGQPPQRRYR 261

Qy 283 RNENYRRRRE-NPKP-ODGKETKA-ADPAENSSAPAEAOGGAE 324

Db 262 RNFNRRROQTGGKPGQESKDAKAGGEPSEAETSAPEAOGGAE 306

Search completed: August 25, 2004, 00:26:20

Job time : 121 secs